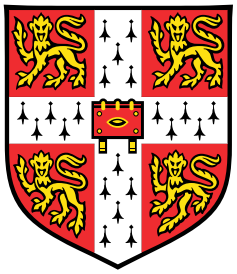
Quantum Telepathy and the Analysis of Particle Presence



David Roland Miran Arvidsson-Shukur

Cavendish Laboratory - Department of Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

King's College

November 2018

This thesis is dedicated to my mother, Ann-Sofie Arvidsson, who introduced me to science, and to my father, Ghazi Shukur, who introduced me to academia.

Declaration

The work presented in this thesis was carried out at the *Thin Film Magnetism* and *Quantum Information* groups at the Cavendish Laboratory, University of Cambridge, between October 2014 and July 2018. I state that this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specified in the text. I further state that no part of my dissertation has already been submitted, or, is being concurrently submitted for any degree, diploma or other qualification at the University of Cambridge or any other University or similar academic institution. The thesis is less than $60,000$ words long and does not exceed the prescribed word limit of the Degree Committee.

David Roland Miran Arvidsson-Shukur
November 2018

Acknowledgements

During the course of my PhD I have enjoyed the support of several individuals whom I want to thank.

I must start by expressing my gratitude towards my supervisor, Professor Crispin H. W. Barnes. Professor Barnes has been a great source of inspiration, academically and personally. His belief in giving students independence and academic freedom has made me particularly privileged. His unparalleled understanding of quantum physics has helped me explore areas far beyond my comfort zone, in the knowledge that a supervising discussion is just a phone call away.

I am grateful to the members of the *Thin Film Magnetism* and the *Quantum Information* groups of the Cavendish Laboratory for providing an intellectually stimulating workplace. In particular I would like to thank: Thierry Ferrus, who has been my secondary supervisor and provided me with a link to industry; Edmund Owen, who patiently trained me during my first years; Jacek Mosakowski, who both taught me "proper" programming and made me aware of the many benefits of PhD life; Eustace Barnes, who has taught me much about the world of nature; Nimalika Perera, who has taught me a great deal about other areas of the natural sciences; and Echo Zhang, Aleksandar Nikolic and Mrittunjoy Guha-Majumdar, with whom I have shared the Cambridge experience.

I acknowledge the funding from the Engineering and Physical Sciences Research Council, and Hitachi Cambridge via the Project for Developing Innovation Systems of the Ministry of Education, Culture, Sports, Science and Technology in Japan.

I have developed valuable and long-lasting friendships during my academic journey. I am grateful to King's College, where a good social life is provided from the first day, and I am deeply thankful towards my friends in *ze Boot*. I want to express particular gratitude towards my close friends: Hippolyte Astier, for standing by my side from the first year of undergraduate studies to the last page of my thesis, showing me good fashion in both academia and life; Hugo Lepage, for working with me on interesting problems in the office, and for being the source of many adventures outside it; Love Hedman, for sharing many of my hopes and ambitions; Axel Gottfries, for our stimulating discussions and our work on the ontology of quantum mechanics; Irati Alonso Calafell, for being a great travel-companion

and an excellent experimentalist; Arthur Duhe, for adding a touch of philosophy to my social life; Reuben Ong, for our many interesting discussions of history; Daniel Chiritoiu, for being the most stimulating, energetic and thoroughly positive friend one could ask for; Bram van der Velden, for being a more realistic friend; Arvid Lundström, for being my oldest friend; Bono Xu, for our many amusing moments and good dinners; Marianne Haroche, for our long conversations and good laughs; Peter Welford, for enriching my non-academic life in Cambridge; Catherine Able-Thomas, for our long-lasting friendship; and, finally, Aleksander Lasek, for being the most loyal of friends—since my first day as an undergraduate—and for being the most valuable of discussion partners in the subject of physics.

Finally, I want to express my thankfulness towards my family. Throughout my life my mother, Ann-Sofie Arvidsson, has introduced me to, taught me and discussed with me the various aspects of culture and science. My father, Professor Ghazi Shukur, has shared with me his deep wisdom in life. Both my parents have been immensely supportive throughout all of my studies. I also want thank my sister, Lydia Arvidsson-Shukur, for making me understand the value of creativity in all aspects of life. Finally, I would like to thank my partner, Eleni Courea, for her wholehearted support, and for always bringing me joy and happiness.

Abstract

The field of quantum mechanics has revolutionised physics as a subject. Areas such as information theory, computer science and physical sensing have all been affected by the tremendous successes of various quantum protocols. In this thesis I present my contribution to the development of such non-classical protocols.

In classical communication theory a message is always carried by physical particles that interact with a transmitter, after which they travel to a receiver. In this thesis I outline a quantum protocol which allows a receiver to obtain a message without receiving any physical object or particles that have interacted with the transmitter—that is, *counterfactually*. I build my protocol for counterfactual communication on the principles of interaction-free measurements, ensuring that information always propagates in the opposite direction to the protocol particles. The protocol shows how quantum mechanics breaks the previous premise of communication theory. From the perspective of local observers, it is a beautiful manifestation of the non-locality of interaction-free measurements. Furthermore, it is highly robust against experimental errors and external disturbances. The majority of this part of the thesis is based on my published article 'Quantum counterfactual communication without a weak trace' [*Phys. Rev. A* 94, 062303 (2016)].

Previous to my work, Salih *et al*. attempted to design a counterfactual communication protocol [*Phys. Rev. Lett.* 110, 170502 (2013)]. This protocol has been highly controversial. As counterfactual phenomena impose restrictions on the inter-measurement paths of quantum particles, and the physical reality of such paths lacks description in the Copenhagen interpretation of quantum mechanics, an extension of current quantum theory is required to facilitate a discussion. In this thesis I present an operational and interpretation-independent methodology, enabling the discussion of inter-measurement paths of quantum particles. I start by considering the interferometers of counterfactual protocols, making the basic assumption that any quantum evolution naturally involves uncontrolled weak interactions. I then show how the Fisher information of these weak interactions, available at the output of counterfactual experiments, can be used to discuss the pre-measurement past of the particles. Based on this analysis, the protocol developed by Salih *et al*. is found to strongly violate counterfactuality. However, my protocol is more flexible in that it allows particles to propagate in the opposite

direction to the message. This leads to counterfactuality being satisfied—even in the presence of large experimental errors. These results are observed both analytically and numerically. This part of the thesis is based on my article 'Evaluation of counterfactuality in counterfactual communication protocols' [*Phys. Rev. A* 96, 062316 (2017)]. The numerical methods are inspired by another of my publications: 'Protocol for fermionic positive-operator-valued measures' [*Phys. Rev. A* 96, 052305 (2017)].

Finally, as the Fisher information measure is found to be useful in evaluating counterfactual protocols, I extend my work by investigating the quantum Fisher information in experiments with general discrete quantum circuits. I prove that the quantum Fisher information of a two-level interaction in a quantum circuit can be expressed by a simple formula. Under certain phase-relations, the formula provides a straightforward connection between the abstract concept of the inter-measurement wavefunction and the quantum Fisher information at the output. With regard to how the information obtained from a certain volume of space influences our perception of classical objects, I argue that the quantum Fisher information measure is highly useful in describing quantum objects. If this measure is applied to observers with a limited set of the experimental measurement outcomes, a quantum object can appear to follow non-classical discontinuous paths. This supports the remarkable conclusion that our perception of the past of a quantum object is subjectively dependent on the measurement we conduct on it.

Contents

List of Figures

Nomenclature

Roman Symbols

$\mathscr{A}$	Quantum observable
a	Eigenvalue of $\hat{A}$
$\boldsymbol{a}$	Optical input vector
$\hat{a}^\dagger$	Bosonic creation operator
$\hat{A}$	Operator corresponding to $\mathscr{A}$
$\hat{a}$	Bosonic annihilation operator
$\langle A \rangle_w$	Weak value corresponding to $\mathscr{A}$
$\boldsymbol{b}$	Optical output vector
$\boldsymbol{c}$	Optical vector before interaction
$\boldsymbol{d}$	Optical vector after interaction
e	Euler's number
$F(\theta)$	Classical Fisher information of the parameter θ
$F_Q(\theta)$	Quantum Fisher information of the parameter θ
g	Coupling constant
H	Horizontal photon state
$\hat{H}$	Hamiltonian operator
$H(\theta : \mathscr{M})$	Shannon mutual information between the parameter θ and the set of measurement outcomes $\mathscr{M}$

i	Unit imaginary number
k	Integer number
$\mathscr{H}$	Hilbert space
$\mathscr{M}$	Set of measurement outcomes
M	Maximum m
m	Integer number
N	Maximum n
n	Integer number
$\mathscr{N}$	Normalisation constant
P	Probability; probability density
p	Momentum variable
r	Reflection coefficient
$\mathscr{S}$	Scattering matrix
t	Time variable; transmission coefficient
$\hat{U}$	Unitary operator
V	Vertical photon state
$\hat{f}$	Unitary two-level rotation
$\hat{V}$	Unitary operator
x	Position variable

Greek Symbols

α	Complex coefficient
β	Complex coefficient
Δ	Spatial width
γ	Time-integrated coupling constant

Λ_θ	Logarithmic derivative with respect to θ
π	Pi $\simeq 3.14\ldots$
Φ	Wavefunction; complex phase
ϕ	Wavefunction; complex phase
Ψ	Wavefunction
ψ	Wavefunction
ρ	Density operator
σ_i	i^{th} Pauli matrix
θ	Unitary operation parameter
ϑ	Unitary operation parameter

Other Symbols

$\uparrow$	Quantum state occupying the upper of two paths
$\hbar$	Reduced Planck's constant
∞	Infinity
$\downarrow$	Quantum state occupying the lower of two paths

Acronyms / Abbreviations

BS	Beam-splitter
CFC	Counterfactual communication
CMZI	Concatenated Mach-Zehnder interferometer
CNMZI	Concatenated nested Mach-Zehnder interferometer
D	Detector
IFM	Interaction-free measurement
MZI	Mach-Zehnder interferometer
NMZI	Nested Mach-Zehnder interferometer

PS	Phase-shifter
QKD	Quantum key distribution
TDSE	Time-dependent Schrödinger equation
TSVF	Two-state vector formalism

Chapter 1

Introduction

1 Quantum Mechanical Motivation

Let me begin by briefly motivating this thesis with respect to the current state of quantum mechanics. Before the discovery of quantum phenomena brought a radical departure from classical physics, it was often obvious how to interpret new theories in relation to the perception of a "physical reality". However, shortly after the beginning of the quantum era, it was clear that a special branch of physics had been found. A large discrepancy was soon revealed between the highly debated—and incomplete—foundations of the field, and the many successes it brought in the development of both theorems and experimental technologies [1–3].

It is remarkable that to this day, there has been little progress in reaching a consensus on a satisfying interpretation of quantum mechanics. The reason is well known: never before have scientists encountered a field of physics where controlled experiments have probabilistic measurement outcomes, and where different experiments on similar systems can result in seemingly contradictory conclusions when analysed with the *classical understanding* of the laws of physics.

Whilst discussing the development of the various interpretations of quantum mechanics is interesting, it is not within the scope of this thesis. Instead, I shall focus on the development and analyses of new quantum protocols. Physicists continue to encounter difficulties in designing a solid ontology[1] of quantum mechanics, and I believe that the ability to do so will require precisely an extension of our current knowledge of quantum phenomena.

With or without a complete understanding of the underlying ontology of quantum mechanics, the development of new quantum protocols requires an operational framework,

[1] Ontology is the study of the nature of reality.

within which we can calculate the probabilistic outcomes of experiments. Even though there has been little progress in creating an ontological formulation of quantum mechanics, there have been tremendous successes in producing a highly accurate framework for calculating the outcome probabilities of quantum experiments. This framework is given by the *postulates of quantum mechanics*.

These postulates provide an immensely successful toolkit, which allows us to calculate measurement outcomes in quantum mechanics [4]. The first attempt to understand the physical nature behind them is known as the Copenhagen interpretation. It was introduced by Werner Heisenberg during his lecture courses in the 1920s, and is based on the work he and Niels Bohr carried out in Copenhagen [5, 6]. Heisenberg and Bohr aimed to find a way to accurately describe the various observations from the then emerging field of quantum mechanics, which from the beginning seemed to defy classical logic.

There are many variations of the postulates of quantum mechanics, but to maintain consistency, I shall use the following formulations in this thesis:

1. The quantum state of a system is described by a generally complex wavefunction, $|\psi\rangle$, which is a vector in the Hilbert space of the system. A single-particle time-dependent spatially distributed quantum state, $|\psi(\boldsymbol{x},t)\rangle$, is subject to the normalisation condition: $\int \langle\psi(\boldsymbol{x},t)|\psi(\boldsymbol{x},t)\rangle\, d\boldsymbol{x} = 1$, where the integral is over all space.[2]

2. Any physical observable, $\mathscr{A}$, of a system has a corresponding Hermitian operator, $\hat{A}$, on the Hilbert space of the system. If a system is described by an eigenvector of the observable, $|\psi\rangle = |a_n\rangle$, a measurement of the observable yields the associated eigenvalue, a_n. The eigenvectors of $\hat{A}$ form a complete set of linearly independent vectors.

3. The immediate reiteration of a measurement of a quantum state will result in identical measurement outcomes.

4. The measurement of an observable, $\mathscr{A}$, of a general quantum state, $|\psi\rangle = \sum \psi_n |a_n\rangle$, yields one of the eigenvalues, a_n, with probability $P_n = |\langle a_n|\psi\rangle|^2$.[3]

5. The expectation value of the measurement of an observable, $\mathscr{A}$, of a quantum state, $|\psi\rangle$, is given by $\langle\hat{A}\rangle \equiv \langle\psi|\hat{A}|\psi\rangle$.

[2] $|\rangle$ represents a quantum state in the familiar Dirac-notation.

[3] This postulate is sometimes referred to as the "collapse postulate" within the Copenhagen interpretation. The allocation of specific probabilities to the measurement outcomes is called the "Born Rule".

6. The evolution of a closed quantum state is a unitary process. The evolution is described by the time-dependent Schrödinger equation: $i\hbar\partial_t |\psi\rangle = \hat{H} |\psi\rangle$, where $\hat{H}$ is the Hamiltonian operator of the system and ∂_t is the partial time-derivative.

These postulates accurately predict the features of current quantum theory.[4]

As with the postulates of quantum mechanics, the probabilistic Copenhagen interpretation has been quoted in several different ways [10–12]. I consider the following summary to be the most appropriate:

1. According to Heisenberg's uncertainty principle, canonically conjugate observables cannot be measured jointly to arbitrary precision.

2. A quantum state is fully represented by its wavefunction without any "hidden" variables.

3. Any observation an experimentalist makes requires a measurement, which is described by classical physics (the limiting case of quantum mechanics). However, the classical terminology of the physical reality is not suitable for describing the nature of quantum states.

4. Between observations, a quantum state evolves according to the unitary evolution of the Schrödinger equation.

5. Upon observation, the quantum state immediately undergoes a non-unitary "collapse" to the corresponding eigenvector: $|\psi\rangle \xrightarrow{\mathscr{A}} |\psi'\rangle \equiv |a_n\rangle$.

6. One should not attempt to ascribe a physical reality to inter-measurement quantum states, but focus on interpreting the observations of the theory. (This has been used to summarise the entire Copenhagen interpretation as "Shut up and calculate!" [13].)

It is clear that the Copenhagen interpretation does not address many fundamental questions—in fact some of the statements even seem contradictory. What triggers the collapse of a wavefunction? What constitutes a measurement? Why are measurements treated differently to other processes?

Physicists have developed various ways of interpreting the results postulated by quantum theory in order to gain a deeper philosophical understanding. The most commonly discussed interpretations are: the de Broglie-Bohm theory [14, 15]; the Many-Worlds interpretation

[4] There exist some suggested counterexamples. For example, the non-linear nature of closed time-like curves cannot be described by standard unitary quantum evolutions [7–9]. However, whether closed time-like curves themselves exist is highly controversial.

[16, 17]; and the Spontaneous-Collapse theory [18, 19]. These are all readily described elsewhere (see, for example, Ref. [20] for a detailed comparison) and I shall only touch on them briefly.

The various interpretations of quantum mechanics distinguish themselves by providing different suggestions about the quantum nature of things—such as determinism, locality and causation. Despite their vast differences in formulation, they all replicate the predictions of the Copenhagen interpretation. Moreover, as they all predict the same outcome probabilities of present-day experiments, it has been impossible to verify the correctness or incorrectness of the different interpretations; and an Occam's razor approach leads back to the Copenhagen interpretation.

Nevertheless, using the known facts about the behaviour of quantum mechanical objects, it is possible to develop and probe new quantum phenomena. John Bell's success in ruling out local hidden-variable interpretations of quantum mechanics, using a beautifully simple inequality, shows us that it is in the development of new quantum processes and theorems that we find hope of one day obtaining a satisfactory description of the quantum world. In this thesis I pursue this search for novel quantum processes.

2 Counterfactuals and Communication

The main goal of this thesis is to develop a protocol for the implementation of a novel quantum phenomenon referred to as "counterfactual communication". By developing a rigid framework for the analysis of such a protocol we shall also find that we can make significant approaches to the fundamental ontology of quantum mechanics. So what is counterfactual communication?

Let me start by introducing the simple and functional definition of the concept of communication used in this thesis: **Communication is a physical process, between two parties (spatially separated by a transmission line), in which one party (the transmitter) sends an informative message to the other party (the receiver).** The way that information has been sent in any previous form of communication is by the *encoding* of a message in physical particles that propagate, directly or indirectly, from the transmitter to the receiver, where the message is *decoded*. This has been the case for protocols that make use of classical particles, as well as for those that make use of quantum particles [21–24]. Figure 1.1 visualises how a message (information) is carried by physical particles in standard communication. Note that throughout this thesis, I shall assume that the operations in the respective "laboratories" of the receiver, the transmission line and the transmitter are independent of one-another. For

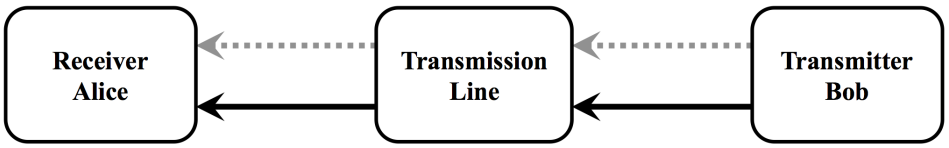


Figure 1.1 The transmitter (Bob) sends a message to the receiver (Alice). In standard communication protocols, information (dotted grey arrow) flows in the direction of the physical particles (solid black arrow).

example, the operations made at the transmitter end cannot affect what operations are applied in the transmission line.

Then, what are counterfactuals? The term was coined by Penrose in 1994 when he discussed the phenomenon of interaction-free measurements (which I shall present in Section 7):

> *Counterfactuals* – that is, things that might have happened, although they did not in fact happen.
>
> – Roger Penrose in 'Shadows of the Mind' [25]

This definition might seem somewhat ambiguous, but it accurately describes the strangeness of a group of peculiar quantum phenomena. The concept will become clearer with the example presented in Section 7 (dedicated to interaction-free measurements), where we shall see how the detection of an object, which classically requires an interaction with the object, can be made counterfactually—in a way such that the interaction *did not in fact happen*. For now it is sufficient to state that counterfactual communication (CFC) is a process in which the transmitter *can* send physical particles to the receiver, but transmits a message *without* this ever happening. In a CFC protocol, the receiver never receives a physical object which has interacted with the transmitter, yet it receives information. That is, CFC would remove the solid black line of particle propagation from the transmitter to the receiver in Fig. 1.1. As a CFC protocol would allow an experimentalist to obtain information from an inaccessible spatial location, without receiving any particles from this location, it has been refered to as "quantum telepathy" in popular science [26, 27].

In order to develop and analyse CFC we shall need to consider a number of different concepts from quantum physics. Moreover, the ability to discuss inter-measurement paths of quantum particles requires arguments to be made in order to extend our understanding beyond the Copenhagen interpretation. In the following sections of this chapter I shall present the relevant theoretical background, which is needed in order to embark on the main objective of the thesis: the development and study of CFC protocols.

3 Quantum Background

3.1 Discrete spatial evolution

The nature of counterfactual phenomena is closely linked to our knowledge about the spatial occupation of the particles that are used in the processes. The Copenhagen interpretation does not provide us with any knowledge about inter-measurement particle *presence*. However, the 6$^{\text{th}}$ postulate tells us how the abstract wavefunction of a particle propagates and spreads in time between measurements. Hence, in order to examine the spatial evolution of a particle's wavefunction one has to solve the Schrödinger equation for a Hamiltonian that adequately represents all the characteristics of the particle, and the potentials of its surroundings. Numerical simulations are often tedious and computationally demanding. A framework for such solutions is, nevertheless, presented in Appendix A. Analytical solutions are generally impossible. However, if all spatial spreading of the individual wavepackets' shape profiles are suppressed such that propagation happens—in a controlled manner—between discrete levels, it is often possible to calculate the evolution of the wavefunction (a complex wavevector) with matrix algebra. This is, for example, the case for photons propagating in optical interferometers. The simplification hinges on the ability to treat a quantum state as point-like or non-dispersive.

Below we study the operational framework of the spatial transitions (called "rotations") between different paths of a quantum circuit. Let us consider the most simple of examples: a quantum circuit made up of only two discrete paths. We denote these paths as "down" and "up", with the respective Dirac notation $|\downarrow\rangle$ and $|\uparrow\rangle$, defined in the following way

$$|\downarrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |\uparrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.1}$$

The unitary matrix that describes the spatial rotation, usually imposed by a beam-splitter, can then be expressed by

$$\hat{U}_{rot} = \begin{pmatrix} r & t \\ -t & r \end{pmatrix}, \tag{1.2}$$

with $|r|^2 + |t|^2 = 1$, where r and t represent the reflection and transmission coefficients, respectively. The phases of the matrix elements are adjustable as long as they satisfy the unitary property, $\hat{U}^\dagger \hat{U} = \hat{1}$.

Consider Fig. 1.2. A quantum wavefunction is input into an interferometer (a quantum circuit for interference) in the lower path: $|\psi_{in}\rangle = |\downarrow\rangle$. The Spartanic interferometer only

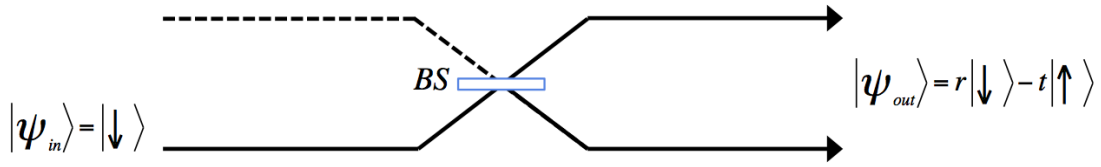


Figure 1.2 The action of a beam-splitter in a quantum interferometer. The discrete particle paths are read as lines travelling from left to right. In this figure, the wavefunction is initially confined to the lower spatial level. After it interacts with the beam-splitter it occupies both the spatial levels.

consists of a single beam-splitter. The action of this beam-splitter (denoted by *BS* in the figure) is to rotate the initial state.

Using the notation above, we can confirm that the action of the interferometer of Fig. 1.2, on a lower path single photon input state, is given by

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \xrightarrow{BS} \begin{pmatrix} r & t \\ -t & r \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} r \\ -t \end{pmatrix}, \tag{1.3}$$

or using Dirac notation:

$$|\downarrow\rangle \xrightarrow{BS} \hat{U}_{rot}\,|\downarrow\rangle = r\,|\downarrow\rangle - t\,|\uparrow\rangle\,. \tag{1.4}$$

It can sometimes be useful to define the rotation matrix in terms of a rotation angle, θ, such that

$$\hat{U}_{rot}(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{1.5}$$

Two rotations by different angles then amount to a combined rotation of the sum of the angles:

$$\hat{U}_{rot}(\theta_1)\hat{U}_{rot}(\theta_2) = \hat{U}_{rot}(\theta_1 + \theta_2). \tag{1.6}$$

The scenario above is of a quantum circuit with only two discrete levels. But a generalisation to circuits with any countable number of levels is straightforward.

3.2 Internal quantum evolution

The nature of any particle is such that it contains an external spatial degree of freedom. However, one must often also consider internal degrees of freedom. In this thesis I shall repeatedly consider particles that possess internal degrees of freedom, such as the polarization of photons or the spin of fermions, which can be independent of the spatial degree of freedom. Let us consider the simple example of a photon bound to a single spatial path, but where

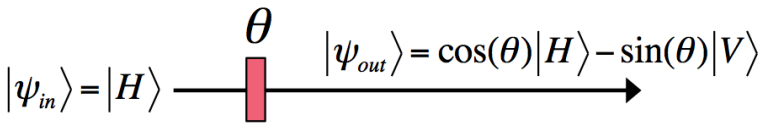


Figure 1.3 The action of a polarization rotator on a photon wavefunction. The interferometer diagrams in this thesis do not show a separate line for the polarization degree of freedom. It is implicit that all spatial photon paths can be occupied by both horizontally polarized and vertically polarized photons.

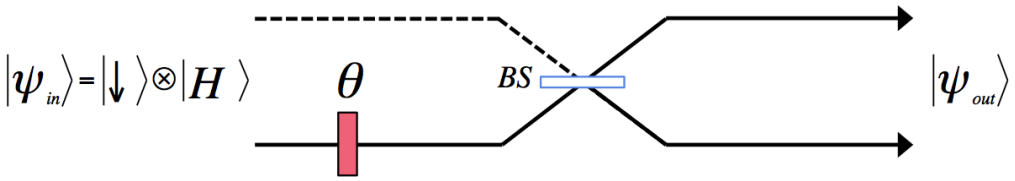


Figure 1.4 A single-photon interferometer consisting of a polarization rotation in the lower path, followed by a beam-splitter joining the lower and upper paths.

the photon also has an internal polarization degree of freedom. Similarly to the spatial basis vectors, the polarization basis can be expressed such that

$$|H\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |V\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.7}$$

The photon can be described by a wavefunction: $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$, where $|\alpha|^2 + |\beta|^2 = 1$. A rotation of the polarization (see Fig. 1.3) can also be represented by Eq. 1.5. The action of the polarization rotator on a horizontal input state is given by: $|H\rangle \rightarrow \hat{U}_{pol}(\theta)|H\rangle = \cos(\theta)|H\rangle - \sin(\theta)|V\rangle$. The extension to rotations of spin-$\frac{1}{2}$ fermion states or photon states with internal angular momentum is straightforward.

For a single particle system without hyper entanglement, internal degrees of freedom are separable from the spatial degrees of freedom. Hence, the outer vector product can be used to mathematically describe the combined wavefunction vector: $|\Psi\rangle = |\psi_{space}\rangle \otimes |\psi_{pol}\rangle$.

Figure 1.4 shows a simple interferometer. First, the polarization of an input state is rotated by an angle, θ. Second, the quantum state is incident on a beam-splitter, which imposes a spatial rotation by an angle, ϑ. Let the input state be a photon of horizontal polarization in the lower arm. The evolution of the wavefunction through the interferometer can then be described by

$$\begin{aligned} |\psi_{in}\rangle = |\downarrow\rangle \otimes |H\rangle \rightarrow |\psi_{out}\rangle =& [\hat{U}_{space}(\vartheta) \otimes \hat{1}_{pol}] \times [\hat{1}_{space} \otimes \hat{U}_{pol}(\theta)] |\downarrow\rangle \otimes |H\rangle \\ =& [\cos(\vartheta)|\downarrow\rangle - \sin(\vartheta)|\uparrow\rangle] \otimes [\cos(\theta)|H\rangle - \sin(\theta)|V\rangle]. \end{aligned} \tag{1.8}$$

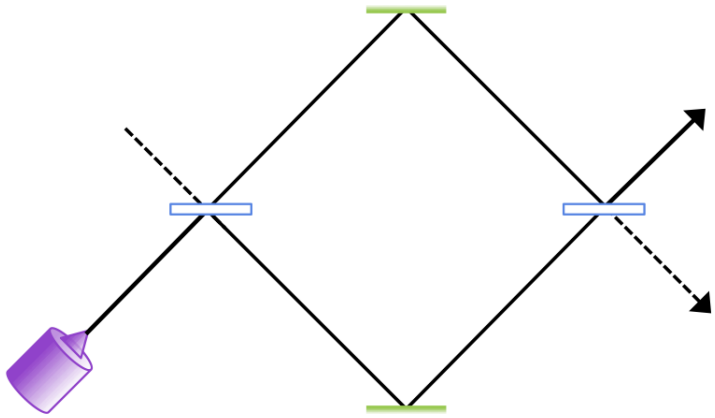


Figure 1.5 An optical diagram of an MZI device.

It is common practice to omit the outer product symbol and the identity matrices in these calculations.

We have now seen the principal workings of the beam-splitters and polarization rotators that will play an essential role in this thesis. Below I shall outline how such objects can be combined to build useful interferometers.

3.3 The Mach-Zehnder interferometer

One of the most commonly used quantum circuits is the Mach-Zehnder interferometer (MZI), developed by Mach and Zehnder in the late 19th century [28, 29]. MZIs are usually considered in an optical framework, but the device can be extended to other systems. The device is simple, it consists of two mirrors and two 50 : 50 beam-splitters arranged in a closed symmetric circuit.

Figure 1.5 shows an optical diagram of an MZI device. The beam-splitters are set such that $r = t = \frac{1}{\sqrt{2}}$, and the mirrors are arranged such that the upper and lower arms of the MZI have equal length. Let us denote a spatial occupation below and above the beam-splitters by $|\downarrow\rangle$ and $|\uparrow\rangle$, respectively, as defined in Eq. 1.1. The unitary evolution of the MZI can then be described by

$$\mathscr{S}_{\text{MZI}} = \hat{U}_{rot}(\pi/4)\hat{U}_{rot}(\pi/4) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{1.9}$$

where $\mathscr{S}$ represents the unitary evolution of an interferometer (often referred to as the scattering matrix). The first beam-splitter transforms an initial state $|\downarrow\rangle$ into a spatial superposition $(|\downarrow\rangle - |\uparrow\rangle)/\sqrt{2}$. The two spatial components then interfere on the second beam-splitter. If the upper and lower interferometer arms are of identical length and free from other optical

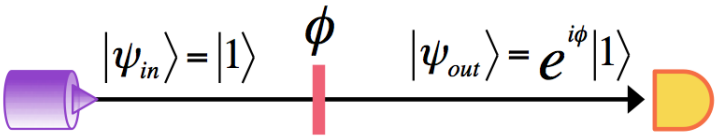


Figure 1.6 A single-photon wavefunction, $|1\rangle$, is given a phase-shift, ϕ, and then measured by a detector.

components, the effect of the second beam-splitter is to create the output state $|\psi_{out}\rangle = -|\uparrow\rangle$. A photon inserted in the lower path will therefore always be output in the upper path, and vice versa.

3.4 Phase measurement

The MZI device outlined in the section above has proven itself to be a remarkable tool, serving as the basic component of many interferometers. The principal mechanism of an MZI is that a single wave is split into two, which travel an equal distance and then interfere. If the two wavefunction components experience identical surroundings we know the exact profile of the output wave. However, if one of the wave components propagates freely whilst the other interacts with a medium, the output distribution will be changed. Let me highlight this with two examples.

Direct phase measurement

Consider an optical object capable of inducing a phase-shift, ϕ, to a photon wavefunction. That is, a wavefunction propagating through the object will evolve such that $|\psi_{in}\rangle \xrightarrow{PS} |\psi_{out}\rangle = e^{i\phi}|\psi\rangle$. The naïve way to measure this phase-shift is to let a photon evolve through the device and then measure it with some sort of detector. Such an attempt is depicted in Fig. 1.6. The probability that the detector measures a single photon is given by $P_1 = |\langle 1|\psi_{out}\rangle|^2 = |\langle 1|e^{i\phi}|1\rangle|^2 = 1$. Unsurprisingly we see that the single photon detector cannot detect the effect of the phase-shift. The phase shift does not affect the single-photon nature of the quantum state in a detectable way; and there simply is no detector that can measure the phase-shift at the end of Fig. 1.6.

Interferometric phase measurement

Let us now, instead, make use of an MZI device. Figure 1.7 shows a setup in which the phase-shifter has been placed in the upper interferometer arm. A single photon is incident on the MZI in the $|\downarrow\rangle$ state. Using the notation from the section above, we can examine the

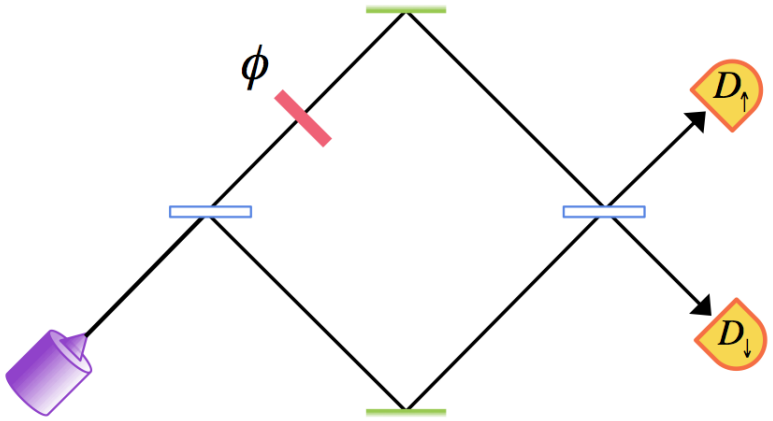


Figure 1.7 An MZI can be used to measure a phase-shift, ϕ.

evolution of the quantum state through the device:

$$|\downarrow\rangle = \begin{pmatrix}1\\0\end{pmatrix} \rightarrow \begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}e^{i\phi} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$$
$$= \begin{pmatrix}\frac{-1+e^{i\phi}}{2}\\ \frac{-1-e^{i\phi}}{2}\end{pmatrix} = \frac{-1+e^{i\phi}}{2}|\downarrow\rangle - \frac{1+e^{i\phi}}{2}|\uparrow\rangle\,. \tag{1.10}$$

The probability of detecting the output state in the upper or lower path is given by $P_\uparrow = \cos^2(\phi/2)$ and $P_\downarrow = \sin^2(\phi/2)$, respectively. We see how the MZI can transfer the information regarding the phase-shift to a spatial degree of freedom. The spatial degree of freedom can then be detected by the single-photon detectors ($D_\uparrow$ and $D_\downarrow$). By performing such an experiment a large number of times, an experimentalist can obtain accurate information about the values of $P_\downarrow$ and $P_\uparrow$. These values can then be used to deduce the value of ϕ. Such a procedure is what the literature refers to as "phase estimation". I shall elaborate on this in the next section.

When defining a wavefunction, the overall phase is often taken to be arbitrary. Just as physicists failed to measure electromagnetic waves with respect to a fundamentally stationary frame of reference (the *ether*), we do not possess a benchmark phase against which we can measure the overall phase of a quantum particle. As we have seen above, we can, nevertheless, measure the phase difference between two components of a wavefunction. Moreover, the example highlights how successful the complex number treatment of the wavefunction is in describing phenomena related to phase-shifts in the theory of quantum mechanics.

In this section we considered a simple example of phase estimation with single-photon quantum states in a two-leveled quantum circuits. However, the extension to multi-photon Fock states and the inclusion of error aspects such as initialisation errors, interferometer

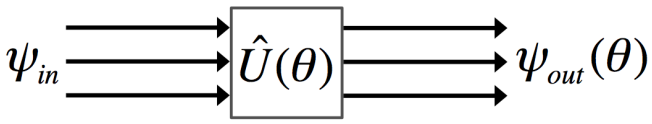


Figure 1.8 A general, known, interferometer (or quantum circuit), with an unknown parameter, θ. The input state is transformed to the output state via the unitary operation $\hat{U}(\theta)$. The transformation depends on θ.

losses and imperfect detections has been readily developed and is presented in, for example, Refs. [30–32].

4 Parameter Estimation

Let us explore further the possibility of estimating unknown parameters with interferometers. We move away from the simplicity of the Mach-Zehnder interferometer, and consider instead a general interferometer. Such an interferometer implements a general unitary evolution, $\hat{U}(\theta)$, which is known up to an unknown parameter, θ. The interferometer acts on a quantum input state, ψ_{in}, and generates an output state, $\psi_{out} = \hat{U}(\theta)\psi_{in}$. See Fig. 1.8. In parameter estimation, the task is to estimate θ from the output wavefunction, ψ_{out}.

Below I outline how experiment outcomes allow us to estimate unknown parameters. The information about a specific value of θ, obtained from the output of interferometers in such experiments, is called the Fisher information (see Chapter 3.3). It is inversely related to the variance of the estimator of θ. In Chapters 3 and 4 we shall study how this information can be used to make inferences about the nature of the underlying quantum state.

4.1 The maximum likelihood method

A common method of estimating an unknown parameter from experimental results is the maximum likelihood method.[5] It is—unsurprisingly—based on the likelihood function, which I shall outline before proceeding to the parameter estimation method.

Consider some realised discrete data (measurement outcomes from an experiment), $x_{1:N} = (x_1, x_2, ..., x_N)$, from a random sample, $X_{1:N} \equiv (X_1, X_2, ..., X_N)$, of sample size N, with independent and identically distributed variables $X_1, X_2, ..., X_N$ (i.e. all variables are mutually exclusive and have the same probability distribution). Observed data is drawn from some distribution with a probability density function given by $P(X = x; \theta)$. The likelihood

[5]The maximum likelihood method does not use any *a priori* information about the unknown parameter. If such information is available various Bayesian estimation techniques can be more suitable [33].

function of the randomly observed data sample can then be defined by the product of the individual probability densities of the observations:

$$\mathscr{L}(\theta|x_{1:N}) := \prod_{n=1}^{N} P(X = x_n; \theta). \tag{1.11}$$

Whilst the statistical concept of *probability* describes the plausibility of the various event outcomes given a parameter, the concept of *likelihood* describes the plausibility of a certain parameter value, given a number of event outcomes [34, 35].

The idea of the maximum likelihood method is to find one or more values of the unknown parameter, θ, that maximises the likelihood function of the observed data $x_{1:N}$. These values, denoted by θ_e, will constitute the estimation of θ. Or more mathematically:

$$\theta_e \in \{\arg\max_{\theta} \mathscr{L}(\theta|x_{1:N})\}. \tag{1.12}$$

Furthermore, the fact that the logarithm is a strictly increasing function can be used in order to simplify the maximisation. It is common to define the log-likelihood function $\ell(\theta|x_{1:N}) \equiv \ln[\mathscr{L}(\theta|x_{1:N})]$, such that

$$\theta_e \in \{\arg\max_{\theta} \ell(\theta|x_{1:N})\}. \tag{1.13}$$

Often, the maximisation can be obtained by explicitly taking the derivative of the log-likelihood function and solving the corresponding equation. Let us proceed with a simple example.

An example: MZI estimation

Again, we start by considering the MZI device from Fig. 1.7. We know that an optical medium will impose a phase-shift to any wavefunction passing through the upper arm. Let this phase-shift be characterised by an unknown parameter, ϕ. In order to estimate the value of ϕ, we evaluate the MZI device by sending through N single-photon states through the lower input path. Say, that we find that detector D_1 ticks n_1 times and D_2 ticks n_2 times, such that $N = n_1 + n_2$. The log-likelihood function is then given by

$$\begin{aligned} \ell(\phi|x_{1:N=n_1+n_2}) &= \ln[\cos^{2n_1}(\phi/2)\sin^{2n_2}(\phi/2)] \\ &= 2n_1 \ln[\cos(\phi/2)] + 2n_2 \ln[\sin(\phi/2)], \end{aligned} \tag{1.14}$$

where the probability densities were calculated in Section 3.4. In order to maximise the log-likelihood function we set the derivative of the function equal to 0 and solve for ϕ:

$$\frac{d}{d\phi}\ell(\phi|x_{1:N=n_1+n_2}) = \frac{2n_2\cot(\phi/2) - 2n_1\tan(\phi/2)}{2} = 0. \tag{1.15}$$

We find that $\ell(\phi|x_{1:N})$ is maximised by $\phi = 2\arctan\left(\sqrt{n_2/n_1}\right)$. If, for example, $n_1 = 75$ and $n_2 = 25$, we have that $\phi = \frac{\pi}{3} + 2\pi k$, where $k \in \mathbb{Z}$. Our estimation is thus given by $\phi_e = \frac{\pi}{3} + 2\pi k$. Note that there is nothing that guarantees that ϕ_e is equal to the exact actual value of the unknown parameter ϕ. However, given the realised data at our disposal ϕ_e is the most likely value of ϕ.

The simple example of phase estimation with an MZI highlights the fact that information of unknown parameters can be encoded in a quantum state during its unitary evolution. In this thesis, I shall not dwell on the details of how to optimise parameter estimation. I shall, instead, consider parameter estimation when examining how information flows through interferometers and between communicating parties in communication protocols. I have already hinted that the Copenhagen interpretation does not provide any means to examine inter-measurement presence of quantum particles. Nevertheless, the main theme of this thesis, that of CFC, hinges upon the knowledge of the inter-measurement whereabouts of quantum particles. In Chapter 3, dedicated to the analysis of counterfactual phenomena, I shall show how the techniques from parameter estimation can be used to make inferences regarding the past of a quantum particle. In Chapter 4 I shall extend and generalise this methodology.

5 The Past of a Quantum Particle via the Weak Trace

At the start of this chapter I suggested that much of the hope of constructing a more rigorous ontology of quantum mechanics is dependent on the development of novel quantum phenomena and theorems. In Section 4, we saw how quantum particles can be subject to interactions between measurements, and that such interactions encode information onto the particles. It thus seems plausible that one could design experiments that utilise these processes so as to improve our ability to discuss the inter-measurement presence of quantum particles.

In 2013 Vaidman outlined a methodology—not directly related to information theory—which attempts to provide information of "the past of a quantum particle" [36]. One should note that Ref. [36] is based on the highly controversial weak trace concept, which has been subject to much debate and criticism (see, for example, Refs. [36–47]). This work, however, beautifully underlines the incentives for developing a new quantum theory from a philosophical perspective.

The ontological implications of the weak trace formalism are often analysed within the framework of the two-state vector formalism (TSVF), which is an alternative interpretation of quantum mechanics. I provide a short summary of the TSVF below.

5.1 The two-state vector formalism

Classical physics is time-symmetric in the sense that that future events can be predicted exactly by previous events, and previous events can be retrodicted exactly by future events. The role of measurements in standard interpretations of quantum mechanics, makes such time-symmetry impossible. If the physical state of a quantum particle at time t is represented by the wavefunction, $\psi(t)$, it only partially predicts the outcome of a complete future measurement, which can probabilistically result in different outcomes.

The two-state vector formalism is a retro-causal interpretation of quantum mechanics that attempts to add time-symmetry to the theory [48]. It predicts the same outcomes from quantum experiments as the Copenhagen interpretation. Nevertheless, it is ontologically different. The "life" of a quantum particle in the Copenhagen interpretation is abstract and physically indeterminate, from the initial observation (pre-selection) of a state $|\psi\rangle$, to the collapse imposed by the final observation (post-selection) of a state $|\phi\rangle$. On the other hand, the TSVF considers the physically real state of a quantum particle at a given inter-measurement time to be represented by a forward evolving pre-selected state, $|\psi\rangle$, *and* a backward evolving post-selected state, $\langle\phi|$. A quantum state subject to a time-independent Hamiltonian, $\hat{H}$, is thus described by $\langle\phi_{out}(t_f)|\, e^{i\hat{H}(t-t_f)/\hbar}$ and $e^{-i\hat{H}(t-t_0)/\hbar}\,|\psi_{in}(t_0)\rangle$ at any time t between the pre- and post-selection. Hence, in the TSVF the present state of a quantum particle is partially described by a future measurement outcome.[6]

5.2 The weak value

We saw earlier how the 4$^{\text{th}}$ postulate of quantum mechanics states that the measurements of identically prepared quantum states result in the probabilistic observations of the eigenvalues of the observables. In the TSVF, however, identical quantum states always yield the same result upon measurement, simply by definition. But what if one is interested in an interaction that takes place between the pre- and post-selection of the particle?

In order to evaluate this question we first outline a framework to treat the process of quantum measurements. Let us temporarily adapt a non-normalised notation and follow the derivation of Duck *et al.* [49]. Consider an interaction that couples the observable operator, $\hat{A}$, of a primary quantum system, $|\Psi_{in}\rangle$, to the position operator $\hat{x}$ of an ancillary quantum

[6] A direct implication of this is that the independence assumption of the Bell inequality is violated.

system (the "pointer"), $|\Phi_{in}\rangle$. The interaction Hamiltonian is given by $\hat{H} = -g(t)\hat{A} \otimes \hat{x}$, where $g(t)$ is the coupling constant of compact support, which is normalised such that its integral over all time is unity. The Hamiltonian generates a unitary evolution operator described by $\hat{U} = \exp(-i\int \hat{H}dt)$, where we have set $\hbar \equiv 1$. The primary quantum system can be decomposed in terms of the eigenvectors, $\{|a_n\rangle\}$, of the observable operator $\hat{A}$:

$$|\Psi_{in}\rangle = \sum_n \alpha_n |a_n\rangle , \tag{1.16}$$

where $\{\alpha_n\}$ are complex coefficients. Let us assume that the position representation of $|\Phi_{in}\rangle$ has a Gaussian profile of width Δ [49]. As momentum is the canonical conjugate variable to position, the momentum representation (given by the Fourier transform) will also have a Gaussian profile, such that

$$|\Phi_{in}\rangle = \begin{cases} \int dx\ \phi_{in}(x)\,|x\rangle , \\ \int dp\ \phi_{in}(p)\,|p\rangle , \end{cases} \tag{1.17}$$

where

$$\phi_{in}(x) = \langle x|\Phi_{in}\rangle = \exp\left(-\frac{x^2}{4\Delta^2}\right), \tag{1.18}$$

$$\phi_{in}(p) = \langle p|\Phi_{in}\rangle = \exp(-\Delta^2 p^2). \tag{1.19}$$

The evolution of the joint quantum system is given by

$$\begin{aligned} |\Psi_{in}\rangle\,|\Phi_{in}\rangle &\to \exp\left(-i\int \hat{H}dt\right)|\Psi_{in}\rangle\,|\Phi_{in}\rangle \\ &= \sum_n \alpha_n \int dx\, e^{ixa_n} \exp\left(-\frac{x^2}{4\Delta^2}\right)|a_n\rangle\,|x\rangle . \end{aligned} \tag{1.20}$$

By using the identity operator $\int dp\,|p\rangle\,\langle p| = \hat{1}$, it is possible to express the output state in the momentum representation:

$$\sum_n \alpha_n \int dp\ \exp(-\Delta^2(p-a_n)^2)\,|a_n\rangle\,|p\rangle . \tag{1.21}$$

Eq. 1.21 shows the output state after the interaction. The relationship between the Gaussian distribution of the pointer and the eigenvectors of $\hat{A}$ suggests that Δ should be large compared to the differences between the eigenvalues, a_n, for the pointer to be able to clearly distinguish between the observable measurement outcomes. If $\Delta \to \infty$ the momentum distribution wavefunction peaks sharply around the eigenvalues. This scenario corresponds to an ideal

"strong" measurement pointer; if a measurement yields a_n, one can be certain that the quantum state is projected onto $|a_n\rangle$. However, if Δ is small compared with the differences between the eigenvalues, the pointer momentum distribution is given by a broad wavefunction. This corresponds to a "weak measurement" [50]; a single weak measurement yields very little information regarding the actual output state of the primary quantum system, as $\Delta p = 1/(2\Delta) \gg \langle \hat{A} \rangle$. After repeated measurements it is, however, possible to determine the expectation value: $\langle \hat{A} \rangle$.

Now, after having studied how an interaction perturbs the joint quantum state of a primary system and a pointer ancilla, we can address the more interesting question regarding an interaction taking place between a pre- and post-selection.[7] Let us now consider a Hamiltonian of the form $\hat{H} = g(t)\hat{A} \otimes \hat{p}$. This time we require that the coupling constant be weak, that is: $\gamma \equiv \int dt g(t) \ll 1$, such that the interaction only weakly perturbs the primary system. We also require that the output state of the primary system is defined (post-selected) as $|\Psi_{out}\rangle$. This scenario is depicted in Fig. 1.9. The final state of the pointer, $|\Phi_{out}\rangle$, can be calculated as follows

$$
\begin{aligned}
|\Phi_{out}\rangle &= \langle \Psi_{out}| \exp(-i\gamma \hat{A}\hat{p}) |\Psi_{in}\rangle |\Phi_{in}\rangle && (1.22)\\
&\approx \langle \Psi_{out}| (1 - i\gamma \hat{A}\hat{p}) |\Psi_{in}\rangle |\Phi_{in}\rangle && (1.23)\\
&= \langle \Psi_{out}|\Psi_{in}\rangle (1 - i\gamma \langle A \rangle_w p) |\Phi_{in}\rangle && (1.24)\\
&\approx \langle \Psi_{out}|\Psi_{in}\rangle \exp(-i\gamma \langle A \rangle_w p) |\Phi_{in}\rangle , && (1.25)
\end{aligned}
$$

where the approximations require that γ is sufficiently small (see Ref. [49]), and where we have introduced the new quantity

$$
\langle A \rangle_w \equiv \langle \Psi_{out}| \hat{A} |\Psi_{in}\rangle \,/\, \langle \Psi_{out}|\Psi_{in}\rangle . \tag{1.26}
$$

$\langle A \rangle_w$ is referred to as the weak value. Remarkably, as the momentum operator generates position translations, the final position representation of the pointer system will be given by

$$
\phi_{out}(x) \approx \phi_{in}(x - \gamma \langle A \rangle_w), \tag{1.27}
$$

and we see that the pointer wavefunction has been spatially translated by a distance corresponding to the weak value. One highly interesting aspect of this post-selected weak measurement is the fact that the value of $\langle A \rangle_w$ may be much larger than $\max_n(a_n)$, the largest eigenvalue of $\hat{A}$. Such anomalous weak values have been investigated in several experiments, see for example Refs. [51–54].

[7] Note that the post-selection corresponds to a strong measurement.

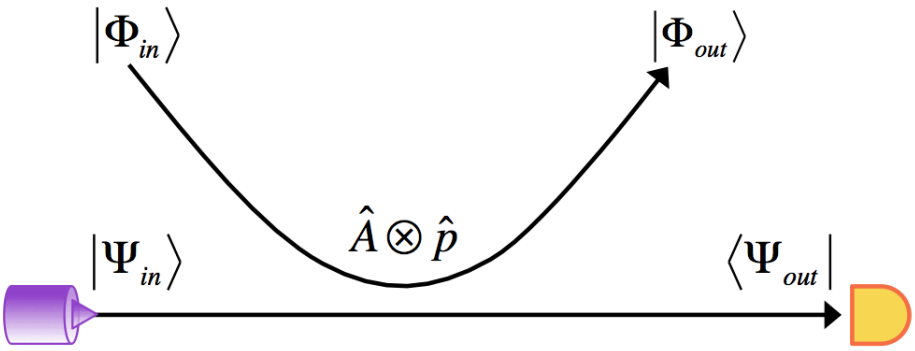


Figure 1.9 A pre-selected quantum state $|\Psi_{in}\rangle$ weakly interacts with a pointer wavefunction, $|\Phi\rangle$, and is post-selected in the state $\langle\Psi_{out}|$.

An example: large weak value

Let us assume that we have access to a quantum pointer capable of making a weak measurement. As the pointer wavefunction plays no direct role in the calculation of the weak value we can neglect its evolution.

Let us also assume that a single-photon source produces a primary quantum state in the form of a polarized photon: $|\Psi_{in}\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$. The simple evolution of the particle is shown in Fig. 1.9. The particle is finally measured, and we require that the post-selection yields $\langle\Psi_{out}| = \cos(\theta)\langle H| + \sin(\theta)\langle V|$. During the evolution of the primary state, it interacts weakly with a pointer state, via the observable operator $\hat{A} = |H\rangle\langle H| - |V\rangle\langle V|$. The eigenvalues of this observable are ± 1. The weak value, however, is given by Eq. 1.26 such that $\langle A\rangle_w = \frac{\cos(\theta)-\sin(\theta)}{\cos(\theta)+\sin(\theta)}$. If $\pi/2 < \theta < \pi$, we obtain weak values larger than the largest eigenvalues of $\hat{A}$ $(+1)$. Actually, if θ approaches $3\pi/4$ from above, the weak value diverges to infinity: $\lim_{\theta\to(3\pi/4)^+}\langle A\rangle_w = \infty$.

With this simple example of a large weak value, we have seen how this post-selected quantity can be a useful tool in the analyses of weak external perturbations experienced by a quantum system at times between initialisation and final measurement. Or in other words, the weak value provides a highly measurable quantity that depends on some inter-measurement interaction.

5.3 The weak trace

In his controversial paper (Ref. [36]), Vaidman argues that the weak measurement framework can be used to map out the inter-measurement "life" (or past path) of a quantum particle. As the pointer wavefunction in a weak measurement is broad enough by definition not to significantly disturb the outcome of an individual experiment, he suggests that the statistical

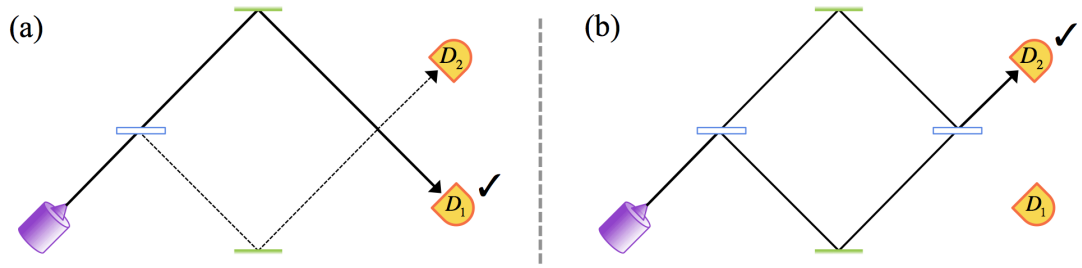


Figure 1.10 An MZI device with the second beam-splitter removed in (a) and in place in (b). A single photon is incident on the first beam-splitter and post-selected on detector D_1 and D_2 in (a) and (b), respectively. The solid line indicates the paths with non-vanishing weak trace.

evaluation of the weak value, due to a weak interaction between measurements, can be used to draw conclusions about where a particle existed between the pre- and post-selection. Experimentally, such a calculation would require measurements of an ensemble of identically pre- and post-selected states—a single evaluation would yield almost no information about the weak value.

Let us consider the examples given in Fig. 1.10. In Fig. 1.10(a) the second beam-splitter of the MZI device has been removed (this can be done at a time after the photon has passed the first beam-splitter). We post-select on detections at detector D_1. The solid curve shows the spatial path where weak measurements would generate non-vanishing weak values. The dashed curve shows the alternative path (if the post-selection had been on detector D_2). In this scenario Ref. [36] concludes that the particle existed in the upper arm of the device. In Fig. 1.10(b), however, the second beam-splitter is inserted such that the quantum self-interference causes the particle to always be detected by detector D_2. In this scenario a non-vanishing weak value is existent in both the upper and lower arm of the MZI. Now Ref. [36] concludes that the particle was present in both these arms simultaneously, and suggests the term "weak trace" for these solid lines representing paths of non-vanishing weak values.

The two scenarios of Fig. 1.10 fit fairly well with the "common sense" picture of quantum mechanics. However, Ref. [36] proceeds with a study of a slightly more complex MZI device, a nested Mach-Zehnder interferometer (NMZI). Such a device is depicted in Fig. 1.11, and is essentially an MZI with another, smaller, MZI nested within it. Whilst the device can be considered in terms of three spatial paths, we follow Ref. [36] and denote the five different spatial locations by A, B, C, D and E. The state $|A\rangle$, for example, means that the wavefunction is localised to the A region, etc.

In Fig. 1.12 we study how the weak trace looks in the NMZI device. Figure 1.12(a) shows the weak trace when we post-select on detection at D_3. To calculate the weak value we label the post-selected wavefunction just before the detection at D_3 as $|D_3\rangle$, and the initial

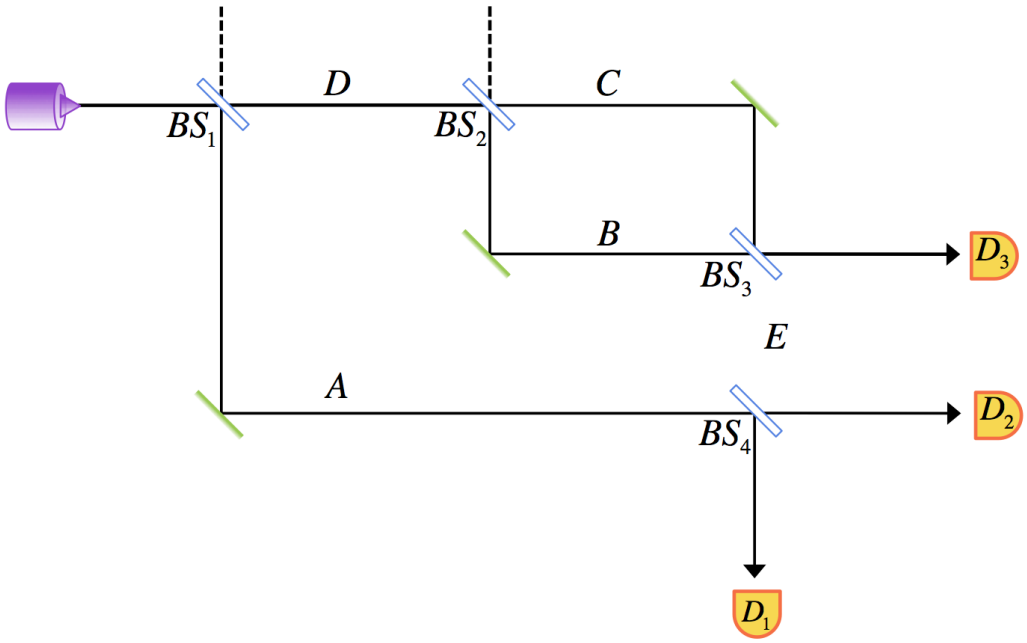


Figure 1.11 A nested Mach-Zehnder interferometer. The beam-splitters' ratios are 50 : 50 and they are tuned such that any wavefunction incident on BS_2 from D will be output from BS_3 onto detector D_3.

wavefunction as $|in\rangle$. The weak value of the operator $|B\rangle\langle B|$ can then be calculated according to Eq. 1.26:

$$\langle B\rangle_w = \frac{\langle D_3|U_{BS_4}U_{BS_3}|B\rangle\langle B|U_{BS_2}U_{BS_1}|in\rangle}{\langle D_3|U_{BS_4}U_{BS_3}U_{BS_2}U_{BS_1}|in\rangle} = \frac{1}{2}. \tag{1.28}$$

Again, it is not too surprising that we find a weak trace at B, C and D, with corresponding weak values $\langle B\rangle_w = \frac{1}{2}$, $\langle C\rangle_w = \frac{1}{2}$ and $\langle D\rangle_w = 1$ given by Eq. 1.26. The weak values at A and E are both 0. However, if we instead post-select on detection at D_2 [see Fig. 1.12(b)] we obtain a more "exotic" result. As expected, we observe a weak value of 0 at D and E. At A the weak value is $\langle A\rangle_w = 1$. The surprising result is that the weak trace is discontinuous; at B and C the weak values are given by $\langle B\rangle_w = \frac{1}{2}$ and $\langle C\rangle_w = -\frac{1}{2}$, respectively.

This discontinuity is even more surprising in the light of the ontological interpretation of the weak trace presented in Ref. [36]. According to the TSVF, the description of a quantum state in an interferometer requires the consideration of both a forward travelling input state and a backward travelling output state. Reference [36] points out that the paths with a non-vanishing weak trace correspond to the paths where the forward evolution of the pre-selected initial state spatially overlaps with the time-reversed evolution of the post-selected final state. This leads the author to conclude that the weak trace corresponds to the paths where the quantum particle has had a physical "past" [36]. According to this interpretation of the weak

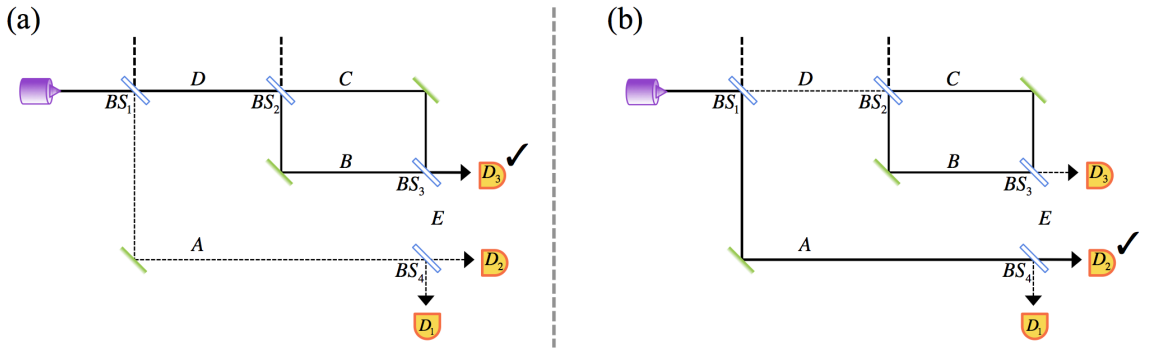


Figure 1.12 The weak trace (solid line) in an NMZI with post-selection at (a) D_3 and (b) D_2, respectively.

trace, a particle detected at D_2 existed within the inner MZI (B and C) although it passed through neither D nor E.

The weak trace evaluation of the past of quantum particles has been subject to intense debate [36–47]. In Chapter 3 I shall show that the reason for the non-vanishing trace within the inner MZI is the slight disturbance of the interference at the third beam-splitter caused by the weak measurement interaction itself. As the measurement of the weak trace slightly perturbs the original interferometer, some authors claim that it cannot be used for a trustworthy evaluation of the evolution within the interferometer [41, 42, 44]. This stance is somewhat incompatible with the essence of quantum theory. Disturbances of the evolution of quantum states are inevitable aspects of quantum theory, ultimately described by the Heisenberg uncertainty principle, as well as an aspect of the unavoidable decoherence in real-world scenarios. I argue that it is impossible to consider the realistic evolution of a quantum particle without any unwanted or uncontrolled interactions. In other words: perfect quantum interference devices are impossible to realise. However, whereas weak traces inside interferometers are experimentally inevitable, I consider the claims of Ref. [36], with respect to the past physical realities of quantum particles, somewhat dubious. These claims seem, at least superficially, to hinge upon the TSVF, which—with its retro-causal time-symmetry—is a niche interpretation of quantum mechanics without the general support of the community. Reference [36] does, nevertheless, provide a crucial piece in the puzzle of establishing a way to think about the past of quantum particles: namely that inter-measurement interactions lead to measurable effects of the output states. In Chapter 3 and 4 I shall use a similar premise of unavoidable weak disturbing interactions and an information-theoretical framework to design an operation and interpretation-independent measure for quantum particle “presence”. This will enable us to truly scrutinise the paths of quantum particles in, for example, counterfactual protocols. But let us now depart from the introduction to inter-measurement measures of

particle presence, and instead look towards another area of quantum physics, which has played an unprecedented role in the development of novel quantum protocols: the quantum Zeno effect.

6 The Quantum Zeno Effect

The quantum Zeno effect can be seen as a consequence of the 3rd and 4th postulates, and it describes how the evolution of a quantum state can be "frozen" if it is measured with high enough frequency [55].

The name of the effect is a reference to the 'Arrow Paradox' presented by the Greek philosopher Zeno of Elea. Zeno stated that at any infinitesimally short instant of time, an arrow only occupies the space where it is located. Zeno then argued that for motion to occur an object has to change its position. Hence, if time is built up of time instances, and if the arrow is stationary at all instances in time, it cannot ever move [56].

Alan Turing is thought to have been the first person to discuss the possibility of a quantum Zeno effect:

> *It is easy to show using standard theory that if a system starts in an eigenstate of some observable, and measurements are made of that observable N times a second, then, even if the state is not a stationary one, the probability that the system will be in the same state after, say, one second, tends to one as N tends to infinity.*
>
> – Robin Gandy recalling Turing's formulation in a letter to Max Newman (Ref. [57])

However, the quantum effect was not properly described by anyone before the works of Degasperis *et al.* [58] and Misra *et al.* [55] in the 1970s. The basis of the quantum Zeno effect can be traced to the postulates of quantum mechanics. Consider an initial quantum state, $|\psi(0)\rangle$, which—for simplicity's sake—is subject to a time-independent Hamiltonian, such that $|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle$. If no intermediate measurements are carried out between time 0 and t, the quantum state will evolve according to the sixth postulate. The probability to measure the state in its initial configuration at time t is thus given by

$$P_0(t) = |\langle\psi(0)|e^{-iHt/\hbar}|\psi(0)\rangle|^2. \tag{1.29}$$

Now, if we consider the probability of remaining in the initial state shortly after $t = 0$, we can expand Eq. 1.29 such that

$$P_0(t) = \langle\psi(0)| \left[1 - \frac{iH}{\hbar}t + \mathscr{O}\left(t^2\right)\right] |\psi(0)\rangle \\ \times \langle\psi(0)| \left[1 + \frac{iH}{\hbar}t + \mathscr{O}\left(t^2\right)\right] |\psi(0)\rangle . \tag{1.30}$$

If the initial measurement of the state, $|\psi(0)\rangle$, is made at $t = 0$, we can calculate the rate of change at that moment:

$$\frac{d}{dt}P_0(t)\Big|_{t=0} = 0. \tag{1.31}$$

This is a crucial result; just after the application of a collapsing measurement the rate of change of a quantum state is vanishingly small. Here we can draw a parallel with the original paradox of Zeno's arrow. In any instant a quantum state is measured, it is momentarily stationary in its quantum evolution. This is the basis of the quantum Zeno effect.

Whilst this result might not seem extraordinary at first, it leads to remarkable results. For example, consider a qubit, in an initial state, $|\psi(t = 0)\rangle = |0\rangle$, subject to a unitary evolution $\hat{U}(t) = \cos(\omega t)\hat{\sigma}_z + \sin(\omega t)\hat{\sigma}_x$. At any point in time after the initialisation at $t = 0$, the qubit state is given by: $|\psi(t)\rangle = \cos(\omega t)|0\rangle + \sin(\omega t)|1\rangle$. For simplicity we take $\omega = 1\,\mathrm{rad\,s^{-1}}$. Let us consider a time-interval of $\pi/2\,\mathrm{s}$. If there are no intermediate measurements between $t = 0\,\mathrm{s}$ and $t = \pi/2\,\mathrm{s}$, the final state is given by: $|\psi(t = \pi/2\,\mathrm{s})\rangle = 1|1\rangle$, and the probability of measuring the initial state is $P_0(t = \pi/2\,\mathrm{s}) = 0$. However, if we instead measure the quantum state every $\pi/40\,s$ a total of 20 times we see that the survival probability, P_0^s, which corresponds to the probability to only observe the initial state, is given by: $P_0^s(t = \pi/2\,\mathrm{s}) = \mathrm{P_0(t = \pi/40\,s)^{20}} = \cos^{40}(\pi/40) \approx 0.88$. By applying 20 quick measurements, it is possible to "interrupt" and heavily suppress the otherwise certain transition of the qubit from the $|0\rangle$ state to the $|1\rangle$ state. A visualisation of the principle behind the quantum Zeno effect is given in Fig. 1.13, where the evolving quantum state described above is measured five times. By increasing the number of intermediate measurements the probability of remaining in the initial state can be taken arbitrarily close to unity.

The examples above give an overview of the basis of the initial development of the quantum Zeno effect. However, the effect is actually not directly bound to the application of measurements within short time-intervals. The necessary requirement is rather that a measurement be carried out after sufficiently small quantum evolutions of the initial state. Consider a quantum particle in the ground-state of a confining potential. Let this particle be perturbed by a weak external potential, applied at a given frequency. It is possible to ensure

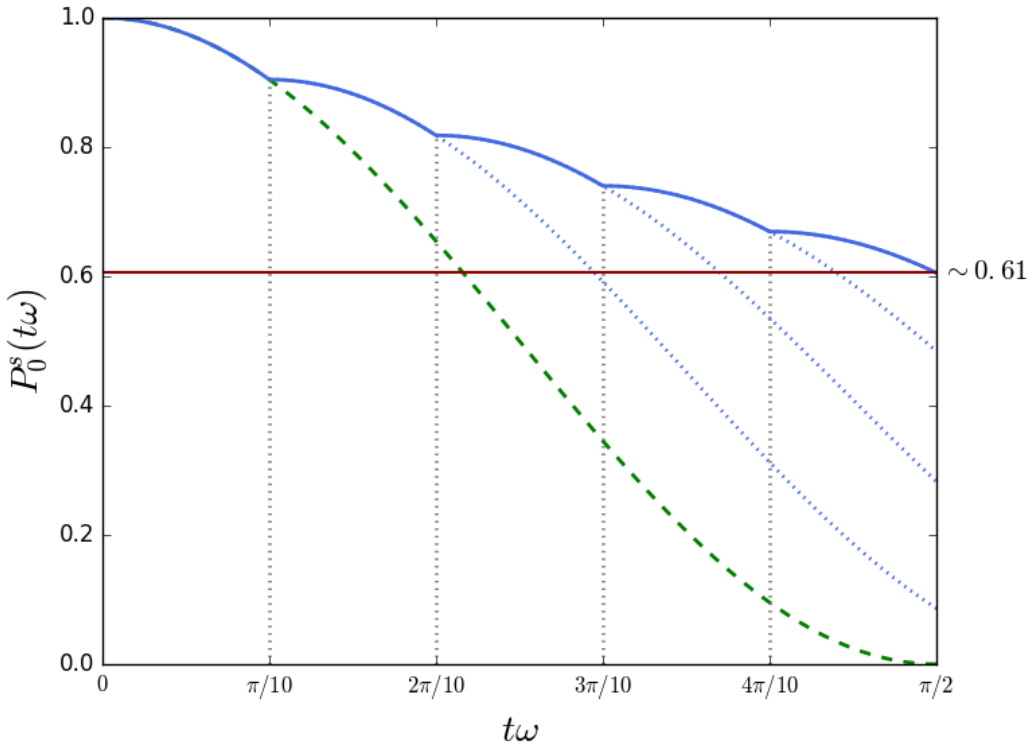


Figure 1.13 An initial quantum state (described in the main text) is measured every $t = \pi/10\,\mathrm{s}$ (dotted vertical lines) from initialisation at $t = 0$ until the final measurement at $t = \pi/2\,\mathrm{s}$. The dashed green curve corresponds to the probability of remaining in the initial state if no intermediate measurements are made. The solid blue curve corresponds to the survival probability of the initial state in the scenario of the intermediate measurements. The final survival probability is given by $P_0^s(\pi/2\,\mathrm{s}) = \cos^{10}(\pi/10) \approx 0.61$.

that the particle stays in the ground state by applying a measurement directly after each—and before the next—perturbation, regardless of the time-intervals between perturbations.

Figure 1.14 shows a single photon state input into a concatenated Mach-Zehnder interferometer (CMZI). The photon is input in the lower state $|\downarrow\rangle$ of the device. However, the interaction with unitary beam-splitters allows the particle to "leak" into the upper path, $|\uparrow\rangle$. In standard CMZI devices, the quantum self-interference allows the wavefunction to build up in the upper path (see Chapter 2.2). Here, however, we consider the scenario of detectors inserted in the upper path, inhibiting the self-interference. If the transmission and reflection coefficients of all the beam-splitters are given by t and r, respectively, the probability of remaining in the lower path after the first beam-splitter is given by $P_{\downarrow}^1 = 1 - |t|^2 = |r|^2$. As detectors are inserted after each beam-splitter, it is easy to convince oneself that the probability to detect the photon in the lower state at the end of the N beam-splitters is: $P_{\downarrow}^N = |r|^{2N}$.

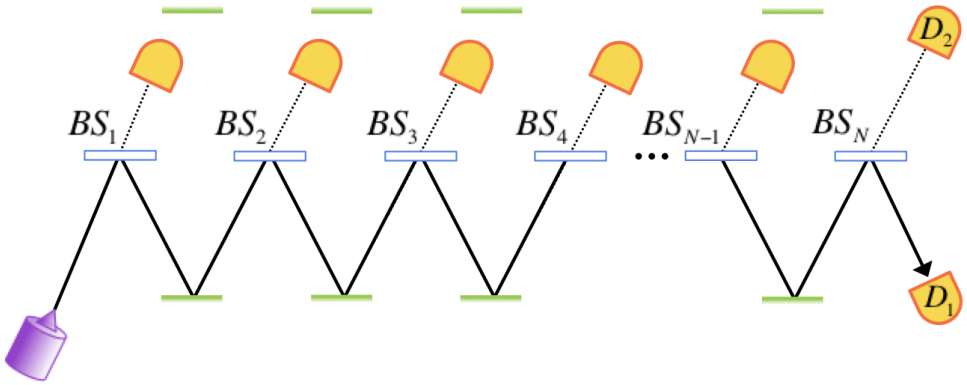


Figure 1.14 Sketch of an optical CMZI. Detectors have been placed in the upper path after each beam-splitter.

In Chapter 2 of this thesis (where CFC is developed) I shall examine beam-splitters that have their reflection and transmission coefficients set with respect to the total number of beam-splitters of the CMZI, such that $r = \cos(\pi/2N)$ and $t = \sin(\pi/2N)$. If no detectors were present, this would cause the single photon state to leave the device from the upper path (see Chapter 2.2). But with detectors inserted we note that $P_{\downarrow}^{N} = \cos^{2N}(\pi/2N)$. An interesting property of this choice of beam-splitter parameters arises in the limit of large N:

$$\lim_{N\to\infty} \cos^{2N}(\pi/2N) = 1. \tag{1.32}$$

Here the transmission coefficient becomes infinitely small, but the device also becomes infinitely long, such that if there were no detectors, there would still be a single photon output in the upper path. However, in the theoretical scenario where there are an infinite number of beam-splitters with transmission coefficients as above, each followed by a detector, Eq. 1.32 implies that it is possible to observe a perfect quantum Zeno effect where the photon always exits from the lower path.

Figure 1.15 shows how the probability of a final photon output in the lower CMZI path of Fig. 1.14 depends on the total number of beam-splitters of the device. Whilst an infinite number of beam-splitters is somewhat unachievable, we are comforted by the fact that for $N \geq 25$ the quantum Zeno effect succeeds with probabilities of over 90 %.

A final point of discussion, which—to my knowledge—has been absent from the existing literature, should be raised with regard to the actual role of the measurement in the quantum Zeno effect. Figure 1.16 shows a modified version of Fig. 1.14, where the detectors as well as the upper mirrors have been removed. In this device, we see that the number of input and output ports have been increased from 1 to N, respectively. A particle, input initially in the

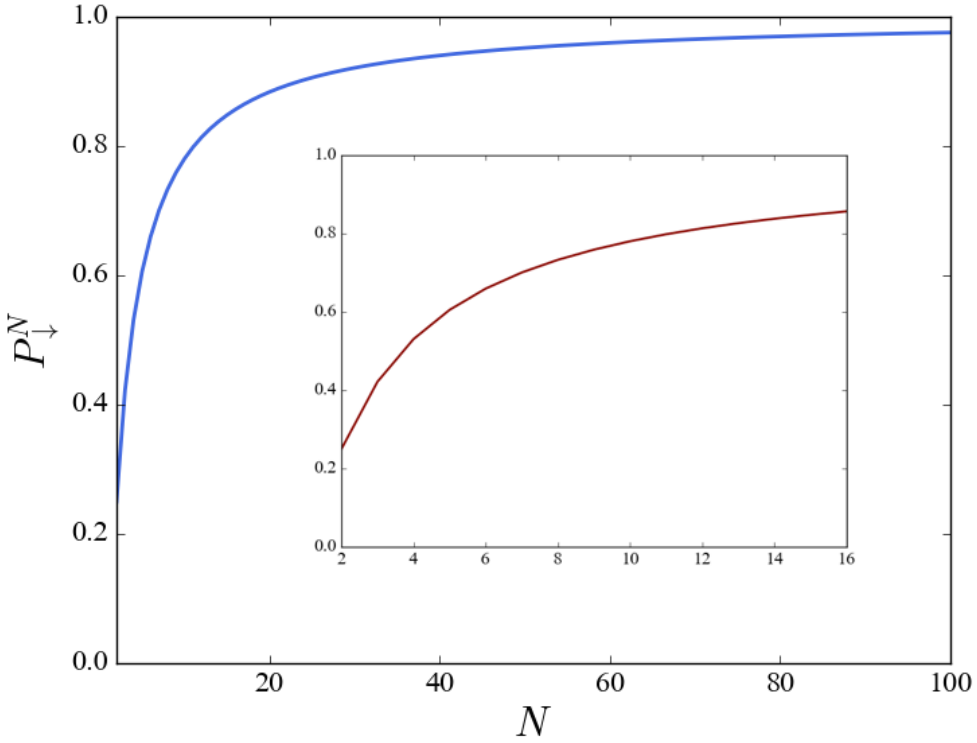


Figure 1.15 The probability of a successful photon output in the lower path of Fig. 1.14, as a function of the number of beam-splitters, N. The reflection coefficients have been set to $\cos(\pi/2N)$.

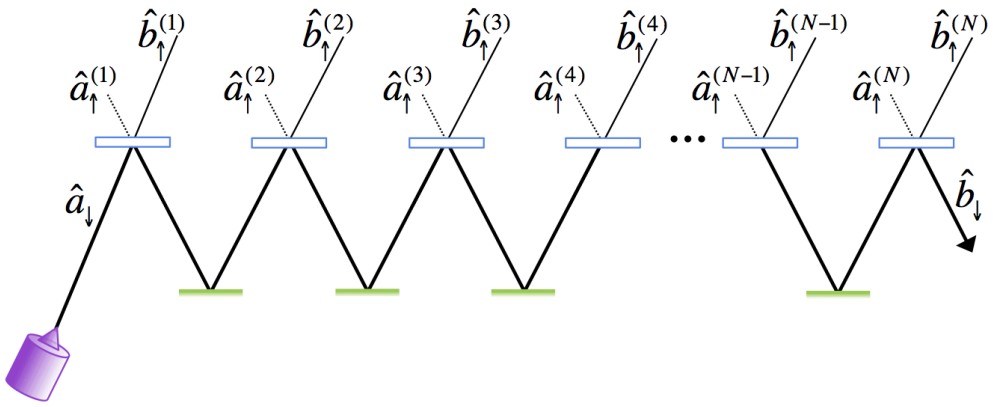


Figure 1.16 The optical device of Fig. 1.14 has been reduced by removing the detectors and upper mirrors. This dramatically increases the spatial Hilbert space of the device.

$|\downarrow\rangle$ state, will evolve such that

$$|\downarrow\rangle \longrightarrow \cos^N(\pi/2N)\,|\downarrow\rangle + \sum_{n=1}^{N} \cos^{n-1}(\pi/2N)\sin(\pi/2N)\,|\uparrow^{(n)}\rangle\,, \tag{1.33}$$

where the state $|\uparrow^{(n)}\rangle$ denotes occupation in the n^{th} upper output slot. (These degrees of freedom can be thought of as constituting an uncontrolled external environment.) Again, we find that the probability of a final output into the lower path is given by $P_{\downarrow}^{N} = \cos^{2N}(\pi/2N)$. In sum, whilst the definition of the quantum Zeno effect requires observations or measurements, there is actually no direct need for intermediate measurements for the manifestation of the effect.

7 Interaction-Free Measurements

We shall soon be in the position to embark on the main project of this thesis: the development and analysis of CFC. But there is one more necessary building block. In Section 4, we saw how an MZI device can be used to estimate unknown parameters. By letting parts of the wavefunction interact with the medium of interest, information is encoded in the quantum state and can, generally, be read (decoded) at the output. Apart from the superposition possibility of having parts of a single quantum particle travelling through different parts of the interferometer, the process is similar to classical parameter estimation. (For example: information about a person's temperature can be encoded in a thermometer via the interaction between the person and the thermometer.) In this section, however, we shall see how the principles of quantum mechanics (in particular the 4^{th} collapse postulate) allow for the ability to obtain information about an object of interest without there being any interactions between the interrogating quantum particle and the object.

The development of the interaction-free measurement (IFM) started with a gedanken experiment. Consider an absurd scenario where you have a number of highly sensitive bombs. Some of these bombs are functioning and some are dud. The task is to remove the dud bombs from the ensemble without destroying the functioning ones. Furthermore, in order to investigate if a bomb is working, it has to be exposed to light. The catch-22 is that if the functioning bombs are inspected they will all detonate, because they are so sensitive that they are triggered by a single photon.

Elitzur and Vaidman were the first to suggest a quantum protocol that partially solved the problem of quantum bomb defusal. In their seminal paper (Ref. [59]) they considered a Mach-Zehnder interferometer, in which a quantum bomb could be placed in the upper

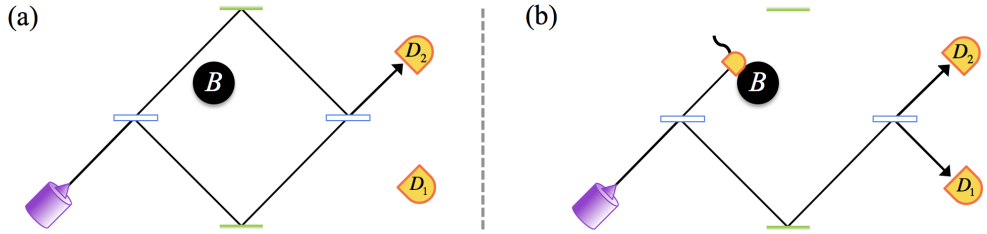


Figure 1.17 An Elitzur-Vaidman bomb tester. An object is (a) absent or (b) present in the upper arm of an MZI device. The output probabilities of detections at D_1 and D_2 are different in the two scenarios. If D_1 is triggered, we can deduce that an object was present in the upper arm, without having interacted with it.

interferometer arm. Figure 1.17(a) shows a scenario in which the bomb is a dud and the single photon detector has been broken off such that the upper arm is empty. In this scenario detector D_2 is always triggered as detections at D_1 are inhibited by destructive interference: $P_2 = 100\,\%$ (see the introduction to MZIs in Section 3.3). On the other hand, in Fig. 1.17(b) the bomb is functioning; the detector is present, and there is therefore an absorbing object in the upper arm. In this scenario, the photon collapses onto the bomb's detector and destroys it with probability $P_{Bomb} = 50\,\%$. However, there is an equal probability that the photon collapses onto the lower path and continues to evolve to the second beam-splitter. In this case there is no destructive interference towards D_1, and the photon is simply split by the second beam-splitter. Hence, the probabilities of detection at D_1 and D_2 are given by $P_1 = 25\,\%$ and $P_2 = 25\,\%$, respectively.

The only way that D_1 can be triggered is if a functioning bomb (an absorbing object) were present in the upper path. Hence, one can find functioning bombs with 25 % probability of success. By reiterating the scheme for inconclusive D_2 detections, the success probability of detecting functioning bombs without detonating them can be made to approach $1/3$. Interestingly—and highly non-classically—the bomb tester protocol sometimes (when D_1 is triggered) allows for the interrogating photon to acquire information about the presence of an object of interest without ever deflecting from or interacting with that object. This is the reason for the term: "interaction-free measurement". Such measurements are examples of the counterfactual phenomena mentioned in Section 2.

Whilst the idea of a single-photon-triggered quantum bomb seems far-fetched, the IFMs have huge potential in the examination of fragile material [60–62]. For example, consider a fragile sample which is to be imaged. Using IFMs the sample could be imagined, pixel by pixel, with a significantly lower probability of photons interacting with, and therefore, destroying it. In Chapter 2 I shall outline how the ideas from the Elitzur-Vaidman protocol

can be extended and used as a building block to enable interaction-free, or counterfactual, communication.

Chapter 2

Counterfactual Communication

Chapter Summary

My main objective in this chapter is to outline two different protocols for the implementation of counterfactual communication (CFC). In Section 2 I present the main optical component of both these protocols, the concatenated Mach-Zehnder interferometer (CMZI). In Section 3, I outline type I CFC—the idea that information can be transmitted between two parties without any particle exchange. I also review the type I protocol published by Salih *et al*. in Ref [39]. This work has been subject to substantial criticism, some of which I summarise in Section 3.2. This criticism was part of the motivation for my own contribution: type II CFC. Type II counterfactuality requires that information be transmitted in the opposite direction to the flow of particles. Section 4 is primarily based on my paper (Ref. [46]) from 2016, published in *Physical Review A*, and co-authored with Crispin Barnes. It contains an outline of type II CFC, as well as a description of the mathematical and optical tools needed to demonstrate such a protocol. The protocol itself is described in Section 4.3. In Section 4.4 I provide a numerical simulation of a Gaussian wavepacket propagating in the type II protocol, confirming that for controlled but realistic quantum evolutions, only minor counterfactual violations arise. Finally, in Section 4.5 I describe a bit-encoding scheme, which makes it possible to significantly reduce the bit error rates of the protocol.

1 Setting the Scene: First Act

It is difficult to overemphasise the enormous impact that the discovery of quantum mechanics has had on physics as a subject. The century-old field has fundamentally changed the way physicists view nature. During this period the physics community has put forward many

new theorems describing the non-classical nature of quantum mechanics. Two well-known examples are Bell's theorem [63] and Hardy's paradox [64]. We have also seen how quantum discoveries led to the possibility of building technologies which solve problems that are not efficiently solvable with classical physics [65–70].

One of the physical novelties that quantum mechanics has provided us with is the interaction-free measurement (IFM) [59]. As we saw in Chapter 1, an IFM is a process in which an interrogating quantum particle is sent through an interferometer and acquires information about whether or not an object is present at a specific location. We also saw that IFMs utilise the 4^{th} postulate in such a way that this information can be acquired without the particle interacting with or deflecting from the object of interest. These interaction-free queries are "counterfactual" phenomena [25].

Unsurprisingly, quantum mechanics has also appended important improvements to classical communication theory. For example, Shannon's classical communication theory [21] from 1948, which outlines how many bits or particles one needs in order to transmit a certain amount of information, had to be changed in order to account for the quantum concept of superdense coding. Whilst the classical theory required that one bit of information be carried by a one-bit particle, the quantum framework, developed by Bennett and Wiesner in 1992 [22], allows the transmitter to encode two classical bits in one quantum-bit (qubit) particle. A further success of quantum communication is described by Schumacher's adaptation of ideas from classical communication theory to the quantum framework, where he showed how classical information can be encoded in qubits and transmitted over quantum channels [23].

Quantum key distribution (QKD) is a particularly famous example of how quantum mechanics improved the limits on secure communication [22, 67, 68, 71–73]. Classically, there can only be unconditional security in communication if the communicating parties have access to the only two copies of a pre-determined one-time-pad. QKD allows the one-time-pad to be created on the go and over a distance. By transmitting bits over a classical channel, in addition to qubits over a quantum channel, a bit-string can be sent securely.

Like QKD, most technologies for quantum information processing have direct analogies in the world of classical physics. One exception is the IFM. See, for example, the section on "bomb defusal" in Chapter 1.7. This is a clear example of where the introduction of quantum phenomena has given us the ability to do something far away from the classical world—in this case the non-destructive monitoring of an object that is destroyed if measured (which in itself constitutes a non-classical concept). Moreover, it has been shown that the ideas from QKD and counterfactual phenomena can be combined such that the one-time-pad can be created without any of the secret key particles travelling between the two communicating

parties [74]. However, no information flows counterfactually in such a protocol, as the protocol also requires transmission of classical bits.

The task of this chapter is to investigate whether quantum physics can make CFC possible. More specifically, is it possible for Bob to transmit a message to Alice telepathically—without Alice receiving any physical object that has interacted with Bob? The motivation for scrutinising the possibility of quantum telepathy is twofold. First, it is immensely important to outline the foundational aspects of communication. The field of communication theory is one of the most important of modern science, and understanding the process of communication at a fundamental level could initiate further key developments. Second, the foundational theory of quantum mechanics is still incomplete, and any prospect of developing new quantum phenomena is appealing in terms of providing additional hints as to how a complete interpretation should be formulated. For example, the aforementioned Bell's theorem [63] has shown us how the behaviour of entangled particles puts constraints on the interpretations of quantum mechanics. The theorem pinpoints the impossibility of a theory based on local realism to reproduce the results of quantum mechanics. CFC would constitute yet another form of quantum non-locality, and might similarly be able to impose further constraints, taking us closer to an accurate interpretation of quantum mechanics.

2 The Concatenated Mach-Zehnder Interferometer

The two protocols for CFC that I present in this chapter are fundamentally different from one another. However, they are both based on a certain optical device, the concatenated Mach-Zehnder interferometer (CMZI). This device has been used by Kwiat *et al.* in order to demonstrate IFMs [75, 76] of significantly higher success rates than the original method proposed by Elitzur and Vaidman [59]. Figure 2.1 shows the CMZI IFM device. The beam-splitters all have transmission and reflection coefficients set to $t = \sin(\pi/2N)$ and $r = \cos(\pi/2N)$, respectively. The device takes single particle inputs in the lower path. The particle's output path is determined by the number of beam-splitters, N, and whether or not there is an absorbing object (e.g. a detector) present in the upper path. For now, let us consider a device with N perfect beam-splitters. If the upper path is left open [Fig. 2.1(a)], the unitary evolution imposed will ensure that the particle be output in the upper path with probability 1. The evolution of the wavefunction is given by

$$|\downarrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xrightarrow{BS_1} \begin{pmatrix} \cos(\pi/2N) & \sin(\pi/2N) \\ -\sin(\pi/2N) & \cos(\pi/2N) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xrightarrow{BS_2} \ldots$$

$$\xrightarrow{BS_N} \begin{pmatrix} \cos(\pi/2N) & \sin(\pi/2N) \\ -\sin(\pi/2N) & \cos(\pi/2N) \end{pmatrix}^N \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -|\uparrow\rangle . \tag{2.1}$$

Here we note that even in the theoretical scenario where $N \to \infty$, such that the beam-splitters have vanishing transmission coefficients, the particle will be transmitted to the upper path. This is a remarkable property of quantum interference.

On the other hand, if the upper path is blocked by detectors [Fig. 2.1(b)], the quantum Zeno effect (see Chapter 1.6) will cause the particle to be output from the lower path with probability $\cos^{2N}(\pi/2N)$. In the limit of large N, this probability approaches unity. To sum up: if an object is not present D_2 always ticks; if an object is present, and N is sufficiently large, D_1 always ticks but D_B and D_2 never ticks. Hence, the success rate of an IFM can be taken arbitrarily close to unity.

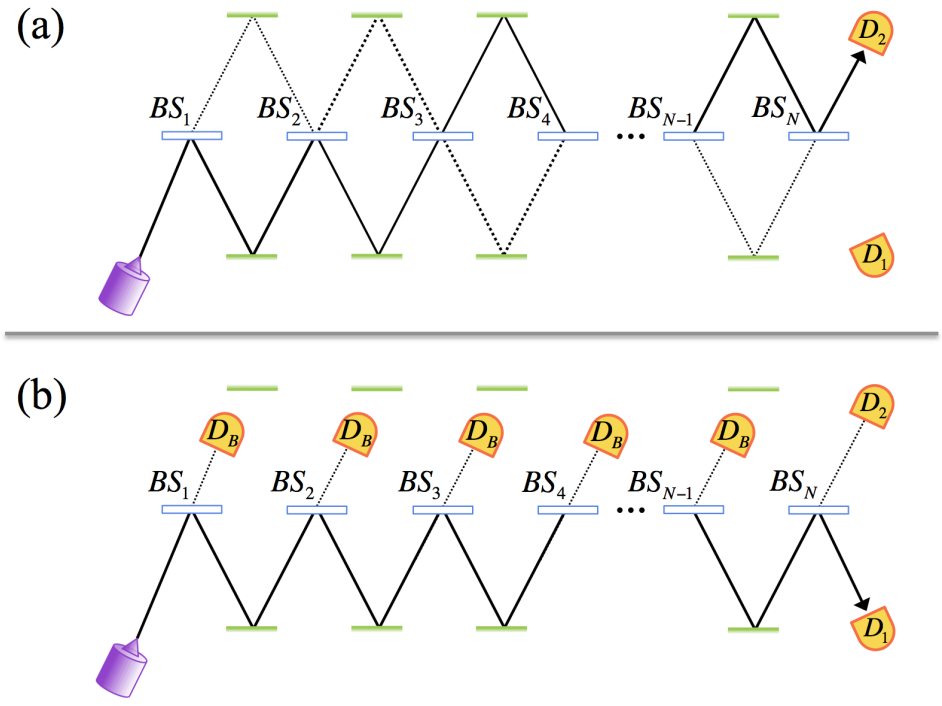


Figure 2.1 The CMZI device with N beam-splitters suggested by Kwiat *et al.* for IFMs. (a) The upper path is left free, such that the particle will evolve and exit from the upper output path. (b) Detectors are inserted in the upper path such that the quantum Zeno effect causes the particle to exit through the lower output path with probability $\cos^{2N}(\pi/2N)$.

Naïvely, one could think that this device could be used for CFC if the lower path was assigned to Alice's laboratory, and the upper path to Bob's laboratory. Indeed, in a scenario

where Bob inserts detectors, Alice would be able to detect this counterfactually. However, if Bob left his upper path open the principles of counterfactuality would be violated. This is because the wavepacket would travel back and forth between the two laboratories before finally being output in Bob's. In order to make the second scenario counterfactual, one would have to eliminate the travels of the wavepacket from Bob's to Alice's laboratory. The differences between the type I and type II protocols for CFC relate to how they address this problem.

3 Type I Counterfactual Communication

Type I and type II CFC are fundamentally different—they have different definitions of the concept of counterfactuality. In Chapter 1 we saw that the term counterfactual was first introduced to describe IFMs—highly non-classical measurements where the interrogating particle obtains information without interacting with the object of interest. The goal of both types of CFC is the same: to send a message from a transmitter, Bob, to a receiver, Alice, without Alice receiving any particles (or waves) that have interacted with Bob.

In this section I will focus on type I CFC. It requires that Alice obtains Bob's message without any particles ever crossing the transmission line. In other words neither Alice nor Bob are allowed to let any of their particles cross the transmission line that separates their respective laboratories. And yet, Alice's particles are supposed to be encoded in accordance with the message that Bob wishes to convey. Such a type I CFC protocol would clearly be non-classical. What is more, CFC would go beyond the previous bounds on quantum communication. Quantum communication has shown how information can be sent more efficiently by the use of quantum channels and quantum particles [22, 23]. However, it has, up to now, not parted from the classical requirement—that particles have to flow in the direction of the transmitted information. Figure 2.2 shows a comparison between classical (Shannon) communication and type I CFC.

3.1 The type I protocol

As outlined in Section 2, a CMZI device can be used to determine whether an object is present in its upper path, without the interrogating particle ever interacting with the object. However, if the object is not present, the interrogating particle (which is input and output in Alice's laboratory) will pass through the spatial location where the object *could* have been present. As this location corresponds to Bob's laboratory, counterfactuality would be violated.

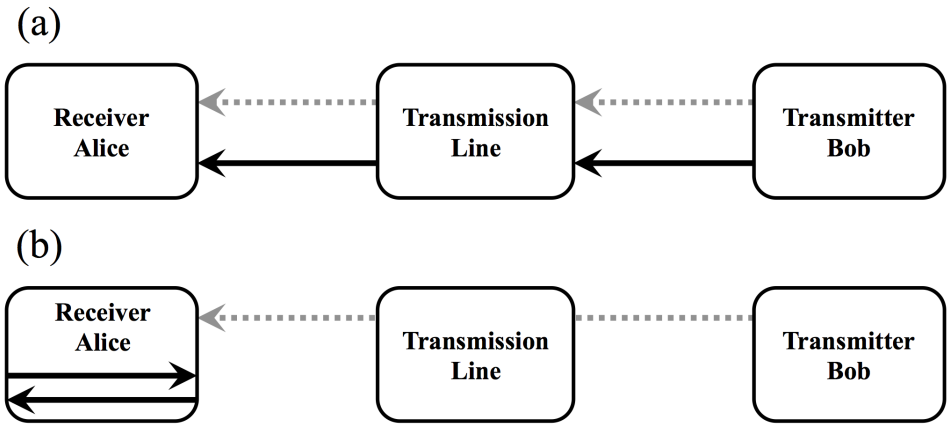


Figure 2.2 The flow of particles (solid black arrows) and the flow of information (dotted grey arrows) in (a) classical communication and (b) type I CFC. Alice is the receiver and Bob is the transmitter of information. The operations within the separate "laboratories" should be independent from one-another.

The first suggestion of how to avoid this for the purpose of CFC, the type I protocol, was presented by Salih *et al.* in 2013 [39].[1] Salih *et al.* base their work on a previous paper by Hosten *et al.* [78], in which the "objects" of the CMZI device are themselves CMZI devices.

If an absorbing object is inserted in the upper path of a CMZI object, the interrogating particle is counterfactually transmitted to the lower output port. If, instead, the upper path is free, the particle will evolve to the upper output path, where it is absorbed by a detector, D_3. Hence, for a CMZI object, the photon is absorbed if there is no absorbing object present in the inner part of the device.

To complicate things further, consider the scenario of investigating the presence of a CMZI object by nesting it inside an outer CMZI device. Figures 2.3 and 2.4 show such a concatenated nested Mach-Zehnder interferometer (CNMZI). The beam-splitters in the inner and outer CMZI device have the reflection coefficient set to $\cos(\pi/2M)$ and $\cos(\pi/2N)$, respectively, where M is the number of beam-splitters in each CMZI object, and N is the number of beam-splitters of the outer CMZI. The protocol requires that $M \gg N \gg 1$. In Fig. 2.3 Bob keeps his laboratory empty, i.e. there is no absorbing object in the upper path of the CMZI object. In this scenario the photon will continuously be collapsed by detectors D_3, and the quantum Zeno effect (see Chapter 1.6) will ensure that it be output on detector D_1. But in Fig. 2.4 Bob is blocking his laboratory, disabling the photon from propagating through the upper path of the inner CMZI. The quantum Zeno effect then causes the photon to be output on detector D_2.

[1] In 2015 I independently developed an equivalent protocol [77].

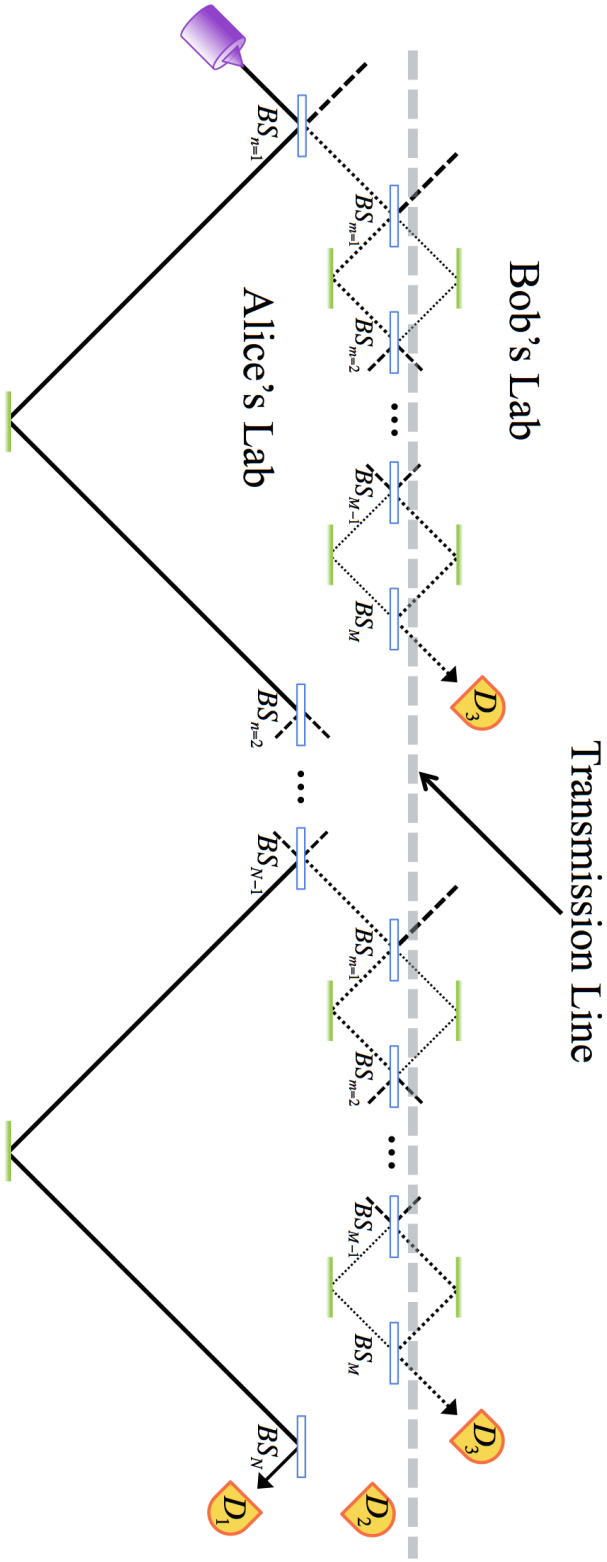


Figure 2.3 The CNMZI device with N outer and M inner beam-splitters used by Salih *et al.* for CFC. If Bob keeps his laboratory open, the photon will be output on detector D_1.

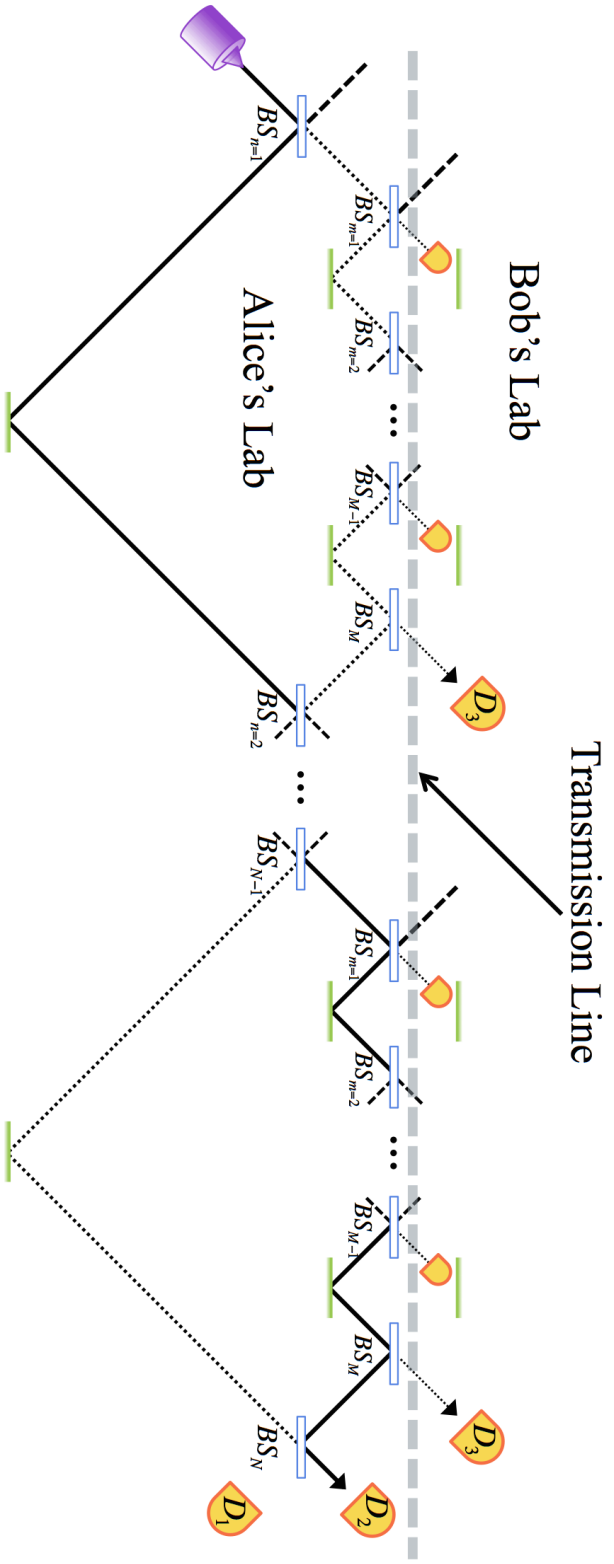


Figure 2.4 The CNMZI device with N outer and M inner beam-splitters used by Salih *et al.* for CFC. If Bob keeps his laboratory blocked, the photon will be output on detector D_2.

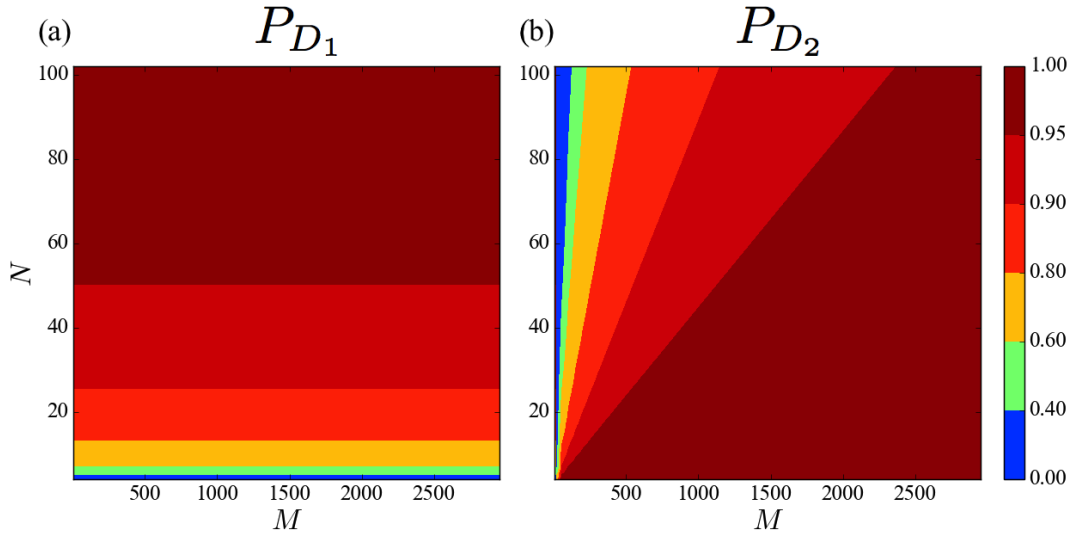


Figure 2.5 The success probability of the (a) logical 0 and (b) logical 1 process of the type I CFC scheme. This figure originally appeared in Ref. [47] and is in agreement with the original paper of Salih *et al.* [39].

The surprising result of this setup is that if $N \to \infty$, $M \gg N$, and the interferometers are perfect, the photon counterfactually ends up at D_2 (D_1) with unit probability of success if Bob does (not) insert detectors in his laboratory. If Bob keeps his laboratory open (blocked) the detection at D_1 (D_2) can be read by Alice as a logical 0 (1). Hence, if Alice and Bob predetermine time-slots at which Alice is to send photons into the CNMZI device, Bob can transmit a message to Alice without any particles ever crossing the transmission line—that is type I counterfactuality. This should hold for both bit-values.

This type I protocol describes some incredible behaviour! So what is the catch? In order to obtain unit probability of success, the number of beam-splitters would have to approach infinity, and their transmission coefficients would approach 0. This is an experimental nightmare. Even if one is prepared to sacrifice some of the success probability, the number of beam-splitters required is huge. Figure 2.5 shows how the success probability of the logical bit processes varies with N and M. A success probability of around 95 % for both bit-values requires that $N \approx 50$ and $M \approx 1,250$ such that the total number of beam-splitters exceeds $60,000$. This should all be worrying for proponents of type I CFC, and as we shall see in the next sections, there are yet further objections to this protocol.

3.2 Criticism of type I counterfactual communication

Since the type I CFC scheme was published, it has been the subject of intense debate and discussion, predominantly by Vaidman [36–45, 47, 79]. This is not surprising given

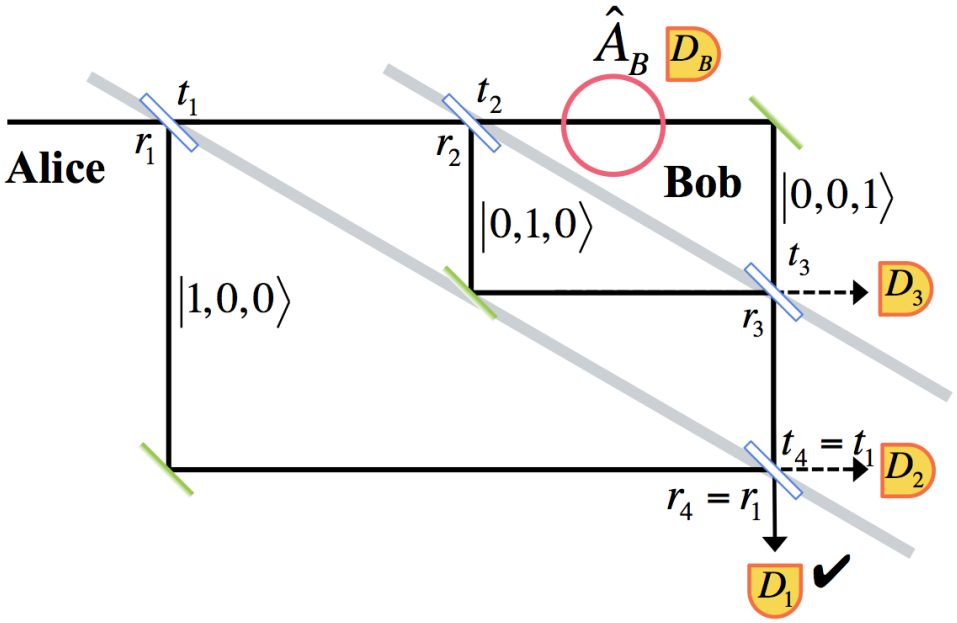


Figure 2.6 The CNMZI device of Fig. 2.3 can be reduced to an NMZI device, which simplifies the conceptual study of counterfactuality in nested interferometers. The two grey diagonal lines separate the spatial occupation states $|1,0,0\rangle$, $|0,1,0\rangle$ and $|0,0,1\rangle$.

the extraordinary claim the protocol makes and the complexity of the device setup. Early criticism of it is based on arguments made with respect to calculations of the weak trace (see Chapter 1.5.3) of the interrogating particle in the CFC scheme.

Given the discussion about weak traces, presented in Chapter 1.5.3 (e.g. that there are no absolutely perfect quantum evolutions), any realistic quantum protocol must allow for minute unwanted interactions to occur, even if they do not significantly alter the end results. We have seen suggestions of how these minute interactions can be used to "trace" the past path of a quantum particle. The very nature of counterfactual protocols is related to the path that the interrogating particle propagates through during the protocol. Thus, if one chose to interpret the weak trace as this past path, it would be a valuable tool to evaluate the "level" of counterfactuality of a process. However, given the vast complexity in the number of optical components exhibited by the CNMZI device, it is a formidable task to map the trace of a quantum particle throughout it. Fortunately, it has been suggested that the conceptual problem of the CNMZI device can be reduced to a study of a single NMZI device [40, 41, 79]. It is with respect to this reduced problem that the weak trace criticism is formalised. Below I summarise the argument as presented by Vaidman in Refs. [40, 79].

First, there is no controversy regarding the counterfactual nature of the logical 1-bit process. This is because if Bob's absorbing object is inserted and the photon is found elsewhere,

it cannot have entered Bob's laboratory.[2] Next, we consider the NMZI device shown in Fig. 2.6. The nested beam-splitters have their transmission and reflection coefficients set to $t_2 = r_2 = t_3 = r_3 = \frac{1}{\sqrt{2}}$. Both outer beam-splitters are identical, with large reflectivity, such that $r_1 = r_4$ and $r_{1,4} \gg t_{1,4}$. Let $|1,0,0\rangle$, $|0,1,0\rangle$ and $|0,0,1\rangle$ represent states spatially confined to the lower left of the outer beam-splitters, in between the outer and inner beam-splitters, and to the upper right of the inner beam-splitters, respectively. The operator that measures presence in Bob's laboratory is thus given by $\hat{A}_B = |0,0,1\rangle\langle 0,0,1|$. Moreover, we post-select on detections by D_1 corresponding to the logical 0 process. The weak value in Bob's laboratory is then given by Eq. 1.26:

$$\langle A_B \rangle_w = \frac{\langle 1,0,0| U_{BS_4} U_{BS_3} \hat{A}_B U_{BS_2} U_{BS_1} |1,0,0\rangle}{\langle 1,0,0| U_{BS_4} U_{BS_3} U_{BS_2} U_{BS_1} |1,0,0\rangle} = \frac{t_1^2}{2r_1^2} \approx \frac{t_1^2}{2}. \tag{2.2}$$

Eq. 2.2 shows how the weak trace is present inside Bob's laboratory during the logical 0 process.[3] For the "weak trace" advocates the conclusion is clear: the interrogating particle existed in Bob's laboratory as it left a weak trace there, and Salih's protocol does not satisfy type I counterfactuality.

Regardless of the opinion on weak traces, the rebuttal by Salih *et al.* contains an ominous remark. They write:

> *Vaidman's argument hinges on the fact that if we measure the weak value of the photon number at C [pink circle in Fig. 2.6], it is nonvanishing. [...] The mistake Vaidman makes is his implicit assumption that any weak measurement in arm C does not affect the interference in the inner interferometer, in direct conflict with the predictions of quantum mechanics.*
>
> – Salih *et al.* in Ref. [41]

Salih *et al.* acknowledge the existence of the weak trace within Bob's laboratory. However, they consider it to be an effect of the weak measurement itself. Moreover, they argue that as a weak measurement is conducted in Bob's laboratory, it is altering the interferometer in a way that destroys its perfect interference, such that it is no longer a study of their protocol.

It is highly problematic for the type I protocol that its counterfactuality is dependent on absolutely perfect quantum channels, free from any disturbances or unwanted interactions. As the number of beam-splitters needed for reasonably high success rates are in the tens of thousands (see Section 3.1), it is highly unrealistic to assume that the protocol can be conducted without the slightest disturbances to the quantum evolution or external interactions.

[2] According to, for example, the Copenhagen interpretation, if D_B is inserted, the wavefunction component that enters Bob's laboratory is erased if the particle is not collapsed and detected on D_B.

[3] See Ref. [79] for a more detailed analysis of the weak trace in the full CNMZI structure.

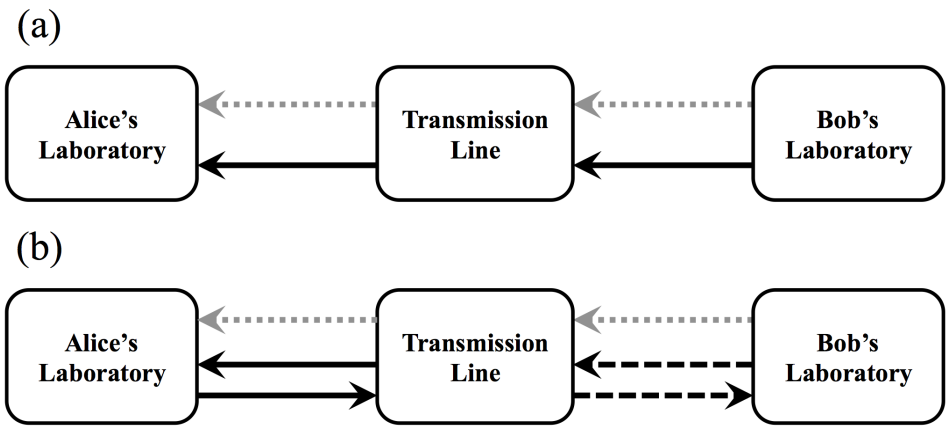


Figure 2.7 Schematic of (a) standard communication protocols and (b) the type II CFC protocol. The dotted grey arrows show the flow of information. (a) The solid black arrows show the flow of particles. (b) The solid black arrows show the flow of particles in the 1-bit process; such particles can enter the transmission line, but will never enter Bob's laboratory. The dashed black arrows show the flow of particles in the 0-bit process; such particles can enter Bob's laboratory and propagate back to the transmission line, but will never re-enter Alice's laboratory. Again, the operations within the separate "laboratories" should be independent from one-another.

In Chapter 3 I shall depart from the weak trace formalism and present a thorough and interpretation-independent study of the counterfactuality of the type I scheme in the presence of unavoidable minute disturbances. It will become even more evident that the type I protocol not only is experimentally unattainable, but also not truly counterfactual.

It was these shortcomings of the type I CFC protocol that motivated me to investigate alternative paths towards realising a truly counterfactual protocol. Is it possible to simplify the device structure? Are the spatial Hilbert spaces of the protocol optimally defined? Could counterfactuality be defined in an alternative way? In the following section I shall approach these questions and outline my own, type II, CFC protocol.

4 Type II Counterfactual Communication

The type II protocol I have developed is based on a new definition of counterfactuality. We must, therefore, begin by considering a few philosophical statements with regard to the "ownership" of quantum particles.

If Alice is capable of manipulating and measuring particles which are localised in her laboratory it is fair to assume that any non-entangled particle existent in her laboratory belongs entirely to her. However, the moment such particles leave her laboratory, and she loses the capability to manipulate or measure them, they should no longer be considered hers. This simple principle has a crucial bearing on our ability to depart from type

I counterfactuality. As we have seen, the type I protocol forbids both Alice's and Bob's particles from crossing the transmission line. However, if Alice's particles were to leave her laboratory and travel to the transmission line, or even to Bob's laboratory, would this necessarily contradict counterfactuality or interaction-freeness? The answer to this question is negative. Remember that the original proposal of counterfactual measurements [59] presented a protocol in which an experimentalist receives information about an object without receiving particles that have interacted with the object. In the type II definition of CFC, a receiver will obtain information from a transmitter without receiving any particles (or waves) that have interacted with the transmitter. However, the transmitter can (and will) receive particles from the receiver. That is, particles in a type II protocol propagate in the opposite direction to the transmitted message. Counterfactuality is retained with respect to the receiver, Alice, by the requirement that none of the particles that left her laboratory and interacted with the transmitter, Bob, are allowed to travel back to her. Figure 2.7 shows a comparison between classical (Shannon) communication and type II CFC. The particles that originate from Alice and travel to Bob (indicated by dashed black arrows) never travel back to Alice. The particles that do travel back to Alice (solid black arrow) never interacted with Bob.

Definition of Type II CFC:
In type II CFC information is encoded in some physical object, which enables the transmitter to communicate to the receiver. If the transmitter interacts with the object, the receiver never receives it (0-bit). If the receiver receives the object, the transmitter has never interacted with it (1-bit).

4.1 The type II device

In Section 2 we saw how a CMZI can be used to carry out an IFM with high success probability. By employing the type II definition of counterfactuality it is possible to define a simple arrangement for Alice's laboratory, the transmission line and Bob's laboratory, which satisfies the counterfactual requirements. Such a device is shown in Fig. 2.8. I will use the following sections to scrutinise this device and describe a protocol in which it can be used for CFC. We shall see that whilst the type I definition leads to complicated device structures and extreme noise sensitivity, the scheme based on the type II definition is realisable with a single MZI device and is counterfactually robust towards device error rates well over 50 %.

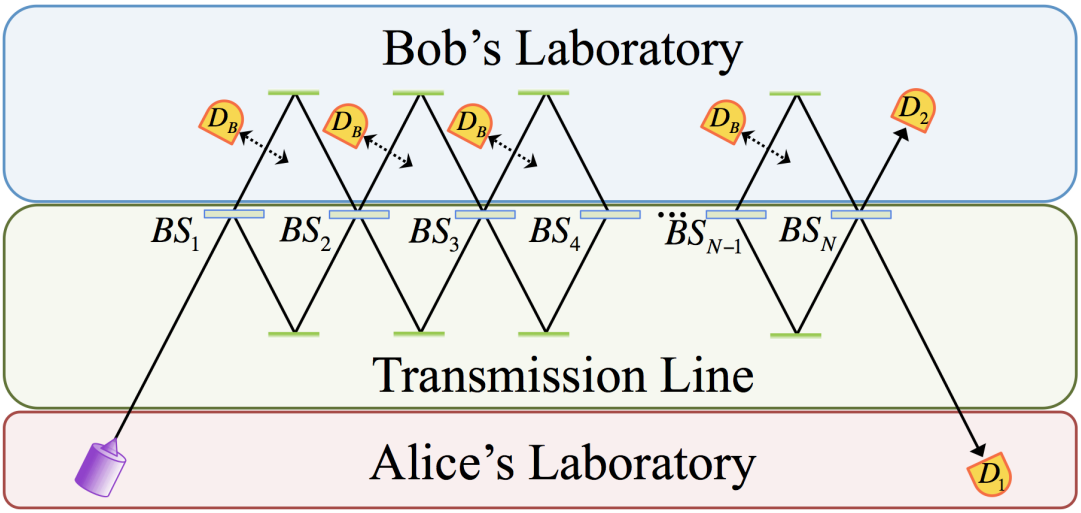


Figure 2.8 The laboratories of Alice and Bob, and the transmission line have been arranged such that the CMZI device of Fig. 2.1 can be used for CFC. Bob's laboratory includes moveable detectors.

4.2 Mathematical framework

In order to develop the workings of the type II counterfactuality scheme we must have a mathematical framework to describe its quantum processes. I shall consider the protocol from a linear-optics perspective and use the formalism of second quantisation. Second quantisation naturally encompasses the particle evolution through the spatial Hilbert spaces, which the successfulness of the protocol depends on.

First, let me introduce bosonic creation and annihilation operators of the respective spatial domains. They are given by $\{\hat{a}^\dagger_{(\mathrm{A})}, \hat{a}^\dagger_{(\mathrm{Tr})}, \hat{a}^\dagger_{(\mathrm{B})}\}$ and $\{\hat{a}_{(\mathrm{A})}, \hat{a}_{(\mathrm{Tr})}, \hat{a}_{(\mathrm{B})}\}$, where, for a general Fock state, $|n\rangle$, $\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle$ for $n > 0$ and $\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle$ for $n \geq 0$.

Second, if we restrict our study to single-photon systems, we can express the basis states by the following vectors

$$
\begin{aligned}
|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} &= |0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}\\
\hat{a}^\dagger_{(\mathrm{A})}|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} &= |1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}\\
\hat{a}^\dagger_{(\mathrm{Tr})}|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} &= |0\rangle^{(\mathrm{A})}|1\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}\\
\hat{a}^\dagger_{(\mathrm{B})}|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} &= |0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|1\rangle^{(\mathrm{B})}.
\end{aligned}
$$

Finally, we need to define the unitary evolutions of the beam-splitters with respect to the annihilation operators. As can be seen in Fig. 2.8, the beam-splitters act between the transmission line and Bob's laboratory. A single beam-splitter implements the following

transformation

$$\begin{pmatrix} \hat{a}'_{(\mathrm{Tr})} \\ \hat{a}'_{(\mathrm{B})} \end{pmatrix} = \begin{pmatrix} i\sin(\theta) & \cos(\theta) \\ \cos(\theta) & i\sin(\theta) \end{pmatrix} \begin{pmatrix} \hat{a}_{(\mathrm{B})} \\ \hat{a}_{(\mathrm{Tr})} \end{pmatrix} = \begin{pmatrix} i\sin(\theta)\hat{a}_{(\mathrm{B})} + \cos(\theta)\hat{a}_{(\mathrm{Tr})} \\ \cos(\theta)\hat{a}_{(\mathrm{B})} + i\sin(\theta)\hat{a}_{(\mathrm{Tr})} \end{pmatrix}. \tag{2.3}$$

where the output and input operators are denoted by primed and un-primed operators, respectively. If we consider the evolution under N such beam-splitters the transformation is given by

$$\begin{pmatrix} \hat{a}'_{(\mathrm{Tr})} \\ \hat{a}'_{(\mathrm{B})} \end{pmatrix} = \begin{pmatrix} i\sin(\theta) & \cos(\theta) \\ \cos(\theta) & i\sin(\theta) \end{pmatrix}^{N} \begin{pmatrix} \hat{a}_{(\mathrm{B})} \\ \hat{a}_{(\mathrm{Tr})} \end{pmatrix} = \begin{pmatrix} i\sin(N\theta) & \cos(N\theta) \\ \cos(N\theta) & i\sin(N\theta) \end{pmatrix} \begin{pmatrix} \hat{a}_{(\mathrm{B})} \\ \hat{a}_{(\mathrm{Tr})} \end{pmatrix}. \tag{2.4}$$

For simplicity's sake I will omit the prime-notation of the operators.

4.3 The protocol

In order to establish counterfactuality it is important to keep track of the spatial occupation of the quantum states of interest. The total Hilbert space of our single-photon system is given by $\mathscr{H} = \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}$, where $\mathscr{H}^{(\mathrm{A})}$, $\mathscr{H}^{(\mathrm{B})}$ and $\mathscr{H}^{(\mathrm{Tr})}$ are the respective Hilbert spaces in the spatial extent of Alice's laboratory, Bob's laboratory and the transmission line. The individual Hilbert spaces have the photon occupation number as their degree of freedom (binary for a single-photon state). Note that the single-photon Hilbert space is a subspace of the multi-photon Hilbert space, allowing for a direct tensor sum rather than a tensor product in the notation above.

Figure 2.8 specifies the spatial extent of the Hilbert spaces with respect to the CMZI IFM device. Alice's Hilbert space, $\mathscr{H}^{(\mathrm{A})}$, contains the lower input path, as well as the lower output path of the device. Bob's Hilbert space, $\mathscr{H}^{(\mathrm{B})}$, contains the upper mirrors of the CMZI, the moveable detectors, as well as the upper output path. Finally, the Hilbert space of the transmission line, $\mathscr{H}^{(\mathrm{Tr})}$, contains the lower CMZI mirrors and all the beam-splitters.

The principle of the communication protocol is simple: Alice is to send single-photon input states into the transmission line of the device at predetermined time-slots. If Alice does not detect a photon at her detector, D_1, within a specific time-slot, she records a 0-bit. If Alice does detect a photon, she records a 1-bit. The two bit processes that make up the CFC protocol are outlined below.

Bit transmission

Step 1: The two (0 and 1) bit processes of the protocol are both initiated by Alice creating a single-photon from the vacuum state:

$$|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} \rightarrow \hat{a}^{\dagger}_{(\mathrm{A})}|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} = |1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}.$$

This step can then be described by $\mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \boldsymbol{\mathscr{H}^{(\mathrm{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}$, where $\mathscr{H}$ and $\mathscr{H}$ denotes an occupied and unoccupied Hilbert space, respectively.

Step 2: Alice lets the photon exit her laboratory and enter the transmission line with the pre-determined frequency:

$$\hat{a}^{\dagger}_{(\mathrm{Tr})}\hat{a}_{(\mathrm{A})}|1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} = |0\rangle^{(\mathrm{A})}|1\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}.$$

The quantum state thus evolves such that $\boldsymbol{\mathscr{H}^{(\mathrm{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})}$.

(a) The 0-bit process:

Step 3_0: Now, if Bob wishes to transmit a 0-bit to Alice, he must make sure that his laboratory is "open", i.e. that the upper path of the CMZI device is free of absorbing objects or detectors. After the single-photon has entered the transmission line it will hit a beam-splitter. This beam-splitter will split the wavepacket of the photon such that one part is reflected and stays in the transmission line, whilst another part is transmitted and enters Bob's laboratory. The angle of the beam-splitters (see Eq. 2.3) is set to $\theta = \pi/2N$. In total, the wavepacket will interfere on N beam-splitters during the evolution through CMZI device. In accordance with Eq. 2.4 the evolution of the quantum state is given by

$$\begin{aligned}\hat{a}^{\dagger}_{(\mathrm{Tr})}|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} &\xrightarrow{BS^N} i\hat{a}^{\dagger}_{(\mathrm{B})}|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})} \\ &= i|0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|1\rangle^{(\mathrm{B})}.\end{aligned}$$

Here, I have omitted the prime-notation as it only describes smaller subspaces of the already defined Hilbert spaces. In terms of the Hilbert space occupation we have that $\mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{B})}}$.

Step 4_0: At this stage of the protocol, the transmission line is emptied into Alice's laboratory (the lower output path of Fig. 2.8 re-enters Alice's laboratory) by the application

of the operator $\hat{a}^{\dagger}_{(\mathrm{A})}\hat{a}_{(\mathrm{Tr})}$. In this process that results in:

$$\hat{a}^{\dagger}_{(\mathrm{A})}\hat{a}_{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{A})}\left|0\right\rangle^{(\mathrm{Tr})}\left|1\right\rangle^{(\mathrm{B})}=0.$$

Step 5_0: In the final step, Alice applies a number operator, $\hat{a}^{\dagger}_{(\mathrm{A})}\hat{a}_{(\mathrm{A})}$, to her state. In the 0-bit process, she will find that her detector does not click and conclude that there is no photon in her spatial domain. She thus records a 0-bit.

A summary of the 0-bit evolution of the quantum state through the Hilbert space is given by

$$\begin{aligned}\mathscr{H}^{(\mathbf{A})}\oplus\mathscr{H}^{(\mathrm{Tr})}\oplus\mathscr{H}^{(\mathrm{B})}\rightarrow\mathscr{H}^{(\mathrm{A})}\oplus\mathscr{H}^{(\mathbf{Tr})}\oplus\mathscr{H}^{(\mathrm{B})}\rightarrow\\ \mathscr{H}^{(\mathrm{A})}\oplus\mathscr{H}^{(\mathbf{Tr})}\oplus\mathscr{H}^{(\mathbf{B})}\rightarrow\mathscr{H}^{(\mathrm{A})}\oplus\mathscr{H}^{(\mathrm{Tr})}\oplus\mathscr{H}^{(\mathbf{B})}.\end{aligned}$$

(b) The 1-bit process:

On the other hand, the 1-bit process of the protocol requires a different procedure to follow Step 2.

Step 3_1: If Bob wishes to transmit a 1-bit, he must ensure that his laboratory is "closed", i.e. that the upper path of the CMZI device has the moveable detectors inserted. According to the 4$^{\text{th}}$ postulate of quantum mechanics the detectors will trigger quantum collapse of the spatial components of the wavefunction that enters Bob's laboratory. This disables the self-interference of the photon on the beam-splitters. Remembering Eq. 1.33, we can outline the evolution of the photon using our second quantisation notation:

$$\begin{aligned}&\hat{a}^{\dagger}_{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{A})}\left|0\right\rangle^{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{B})}\\ &\xrightarrow{BS_1}\left[\cos(\theta)\hat{a}^{\dagger}_{(\mathrm{Tr})}+i\sin(\theta)\hat{a}^{\dagger}_{(\mathrm{B})}\right]\left|0\right\rangle^{(\mathrm{A})}\left|0\right\rangle^{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{B})}\\ &=\cos(\theta)\left|0\right\rangle^{(\mathrm{A})}\left|1\right\rangle^{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{B})}+i\sin(\theta)\left|0\right\rangle^{(\mathrm{A})}\left|0\right\rangle^{(\mathrm{Tr})}\left|1\right\rangle^{(\mathrm{B})}\\ &\rightarrow\begin{cases}\left|0\right\rangle^{(\mathrm{A})}\left|1\right\rangle^{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{B})}, & \text{with } P=\cos^2(\theta),\\ \left|0\right\rangle^{(\mathrm{A})}\left|0\right\rangle^{(\mathrm{Tr})}\left|1\right\rangle^{(\mathrm{B})}\rightarrow collapse, & \text{otherwise.}\end{cases}\\ &\xrightarrow{BS_2}\ldots\\ &\xrightarrow{BS_N}\begin{cases}\left|0\right\rangle^{(\mathrm{A})}\left|1\right\rangle^{(\mathrm{Tr})}\left|0\right\rangle^{(\mathrm{B})}, & \text{with } P=\cos^{2N}(\theta),\\ \left|0\right\rangle^{(\mathrm{A})}\left|0\right\rangle^{(\mathrm{Tr})}\left|1\right\rangle^{(\mathrm{B})}\rightarrow collapse, & \text{otherwise.}\end{cases}\end{aligned}$$

In terms of the Hilbert space occupation, we note that with probability $\cos^{2N}(\theta)$, a successful evolution can be described by $\mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})}$. An unsuccessful evolution happens with probability $1-\cos^{2N}(\theta)$ and should instead be described by $\mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{B})}} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}$.

Step 4_1: The protocol again empties the transmission line into Alice's laboratory. This time the evolution is given by

$$\begin{aligned}\hat{a}^{\dagger}_{(\mathrm{A})}\hat{a}_{(\mathrm{Tr})} &\begin{cases} |0\rangle^{(\mathrm{A})}|1\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}, & \text{with } P=\cos^{2N}(\theta), \\ collapse, & \text{otherwise.}\end{cases} \\ &= \begin{cases} |1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}, & \text{with } P=\cos^{2N}(\theta), \\ collapse, & \text{otherwise.}\end{cases}\end{aligned}$$

With probability $\cos^{2N}(\theta)$, we have that $\mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \boldsymbol{\mathscr{H}^{(\mathrm{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}$. And with probability $1-\cos^{2N}(\theta)$ we have that $\mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}$.

Step 5_1: At this point Alice applies the number operator, $\hat{a}^{\dagger}_{(\mathrm{A})}\hat{a}_{(\mathrm{A})}$, to her state:

$$\begin{aligned}\hat{a}^{\dagger}_{(\mathrm{A})}\hat{a}_{(\mathrm{A})} &\begin{cases} |1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}, & \text{with } P=\cos^{2N}(\theta), \\ collapse, & \text{otherwise.}\end{cases} \\ &= \begin{cases} 1\,|1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}, & \text{with } P=\cos^{2N}(\theta), \\ 0, & \text{otherwise.}\end{cases}\end{aligned} \tag{2.5}$$

With probability $\cos^{2N}(\theta)$ the process succeeds and Alice detects a particle in her laboratory. In this scenario she records a 1-bit.

We can thus summarise the quantum evolution through the Hilbert space of the successful 1-bit process:

$$\begin{aligned}\boldsymbol{\mathscr{H}^{(\mathrm{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \\ \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \boldsymbol{\mathscr{H}^{(\mathrm{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}.\end{aligned}$$

Hence, in the successful scenario, the wavepacket never enters Bob's laboratory and Alice received the 1-bit counterfactually. The unsuccessful evolution can be described by

$$\boldsymbol{\mathscr{H}^{(\mathrm{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow$$

$$\mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \boldsymbol{\mathscr{H}^{(\mathrm{B})}} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}.$$

In the unsuccessful scenario the wavepacket travels from Alice to Bob but never vice versa. This would generate a bit error, but would still be type II counterfactual (as no wavefunction components propagate from Bob to Alice).

Remember that in the limit of an infinite number of beam-splitters:

$$\lim_{N\rightarrow\infty} \cos^{2N}(\theta = \pi/2N) = 1.$$

This means that if $N \rightarrow \infty$, and the beam-splitters are perfectly tuned, the success probability of the 1-bit process approaches unity—an optical manifestation of the quantum Zeno effect [55, 58, 76]. Bob suppresses the evolution into his Hilbert space by his frequent measurements and collapse of infinitesimal parts of the photon wavefunction.

In this way, we have provided a protocol containing two different processes, which allow Bob to send a 0-bit or a 1-bit to Alice. From the analysis above we can see that the message (from Bob) and the flow of particles (from Alice) are counterpropagating. With reference to Fig. 2.7 we conclude that our protocol satisfies type II counterfactuality. The ability to satisfy the definition of type II CFC hinges on the self-interference ability of a single quantum particle. The transmission line conducts the same operation during both bit processes, independently of Bob's actions, yet still enables the encoding of two different bits in the single-particle object.[4] Owing to its apparent similarity to various mythological accounts of telepathy, the type II protocol has been referred to as "quantum telepathy" in the media [26, 27].

However, despite my earlier promise, the type II protocol still seems to require a large number of optical components for high success rates. Below I shall show that this can be avoided. However, before that, I shall present a brief analysis of the possibility of counterfactual violations when the protocol is considered from a more realistic perspective—without point-like particles and perfect quantum operations.

[4] It might appear as if the protocol above could be implemented with classical waves (such as classical light waves). However, from a quantum perspective, a "classical" light wave incident on a beam-splitter would always lead to the transmission of a finite number of particles, which subsequently interact with Bob in both bit processes. In such a scheme the entire wave makes up the object in which information is encoded, and the definition of type II CFC (given in the beginning of this section) would thus not be satisfied for the 1-bit process.

4.4 Numerical demonstration

In the section above we used the mathematical framework of Dirac notation and second quantisation to examine the quantum evolutions that successfully implement the type II CFC protocol. There are, however, a couple of remarks to be made with respect to this framework. First, we should not forget the importance of the 6^{th} postulate of quantum mechanics, which states that the unitary evolution of a wavefunction is given by the Schrödinger equation, and not by some point-like propagation. In the section above we completely neglected the spatial profile of the wavefunction and represented the unitary operations with heavily discretised matrix transformations. For many optical circuits this is a reasonable simplification—but it is never exact. By neglecting the spatial extent of the wavefunction components we lose knowledge of how an actual wavefunction would propagate between the optical devices (mirrors and beam-splitters). In order to account for these short-comings I shall evaluate the time-dependent Schrödinger equation (TDSE) of a Hamiltonian that implements the type II CFC protocol.

In this thesis I use a numerical staggered leapfrog method to solve the Schrödinger equation. See Appendix A for an outline of this algorithm. Here I shall use it to answer the question: To what extent does the type II protocol satisfy counterfactuality if we consider a finite size quantum particle propagating through the spatial Hilbert space?

Let us approach this question by investigating the numerical solution of the TDSE for a charged massive single-particle Hamiltonian, tailored to implement the CFC protocol from Section 4.3. Instead of considering point-like photons, we now use a massive Gaussian particle as a toy-model. For simplicity's sake, let us require that the particle be confined in two dimensions and free to propagate in only the x-direction. Moreover, we suppress the spatial dispersion of the Gaussian wavepackets by placing the wavefunction components in harmonic potentials. Spatial translations of the individual wavefunction components can then be realised by imposing diabatic shifts to the loci of the minima of the potentials. If these shifts are implemented at times when the wavefunction is near-stationary, non-dispersive spatial translations can be realised. Whilst our toy-model is heavily simplified and far from a realistic experimental scenario, it does provide useful information about the spatial wavefunction evolution throughout the CFC protocol. Moreover, the individual components of the protocol are no more complex than those considered in Ref. [80], which implements a Fermionic positive-operator-valued measure [81] within a surface acoustic wave framework [82]. Therefore, in principle, a more elaborate numerical model could be used to represent a more realistic experimental scenario. The spread of the wavefunction components in such a model would, however, be near-identical to results from the simple model, and not yield

more information regarding the counterfactuality of the quantum evolution. For these reasons we can stick with the toy-model.

Figure 2.9 shows the evolution of the wavefunction through the type II protocol with $N = 7$. The figure shows the protocol's successive time-frames (top to bottom). The Hilbert space occupation is, again, represented by bold font. A weak occupation, only caused by small protocol errors, is represented by a prime.

An electrically charged Gaussian particle is initialised in Alice's laboratory (A). As the particle is situated to the left of the minimum of a harmonic potential, it will propagate coherently towards the right and enter the transmission line (Tr). When the particle has reached the top of the harmonic potential and momentarily is stationary in the transmission line (at $x = 2.5$), the potential is given a diabatic shift. This causes the wavepacket to again travel to the right and hit the beam-splitter situated at the bottom of the harmonic potential (i.e. at $x = 3.5$). One part of the wavefunction will rebound into the transmission line and another part will enter Bob's laboratory (B). The harmonic potentials are kept static for the time needed for the particle to interfere on the beam-splitter seven times. (The time-frames of the $n = 2$ to $n = 6$ interactions are not shown in the figure.) After the seven interactions, another diabatic shift is made to the harmonic potentials, causing the content of the transmission line to be emptied into Alice's laboratory.

In Fig. 2.9(a) Bob keeps his laboratory free of detectors, whilst he in Fig. 2.9(b) imposes collapse at $x = 4.5$. In Fig 2.9(a) we can see that the constructive interference on the beam-splitter results in a transfer of the wavepacket to Bob's laboratory (in accordance with the protocol outlined in the section above). Note that in Fig. 2.9(a), the vertical axis scaling of Alice's laboratory and the transmission line in the last two frames has been magnified such that a small failure probability density is visible. This small part of the wavefunction, which generates 0-bit errors, is an effect of the creation of excitations of the Gaussian wavepacket throughout the protocol, which affects the fidelity of the beam-splitters. In this simulation the errors generate a 0-bit failure probability of $\sim 0.95\,\%$, not considered in the initial outline of the protocol. Moreover, following the wavefunction propagation through Fig. 2.9, it is clear that such errors result in rare violations of counterfactuality.[5]

Figure 2.9(b) shows the scenario where Bob inserts detectors in his laboratory but fails to detect the particle seven times. As predicted by our protocol, this results in Alice detecting a particle that has never existed in Bob's laboratory. The success probability of this is $\sim 70\,\%$, in accordance with Eq. 2.5. The figure also confirms that an unsuccessful 1-bit event will not

[5] We note that if the protocol was evaluated with a classical wave, a slight disturbance in the interferometer would cause a finite part of the wave to always end up at Alice's laboratory, such that counterfactuality would be always violated. It is the ability of single-particle self-interference that allows us to overcome this issue.

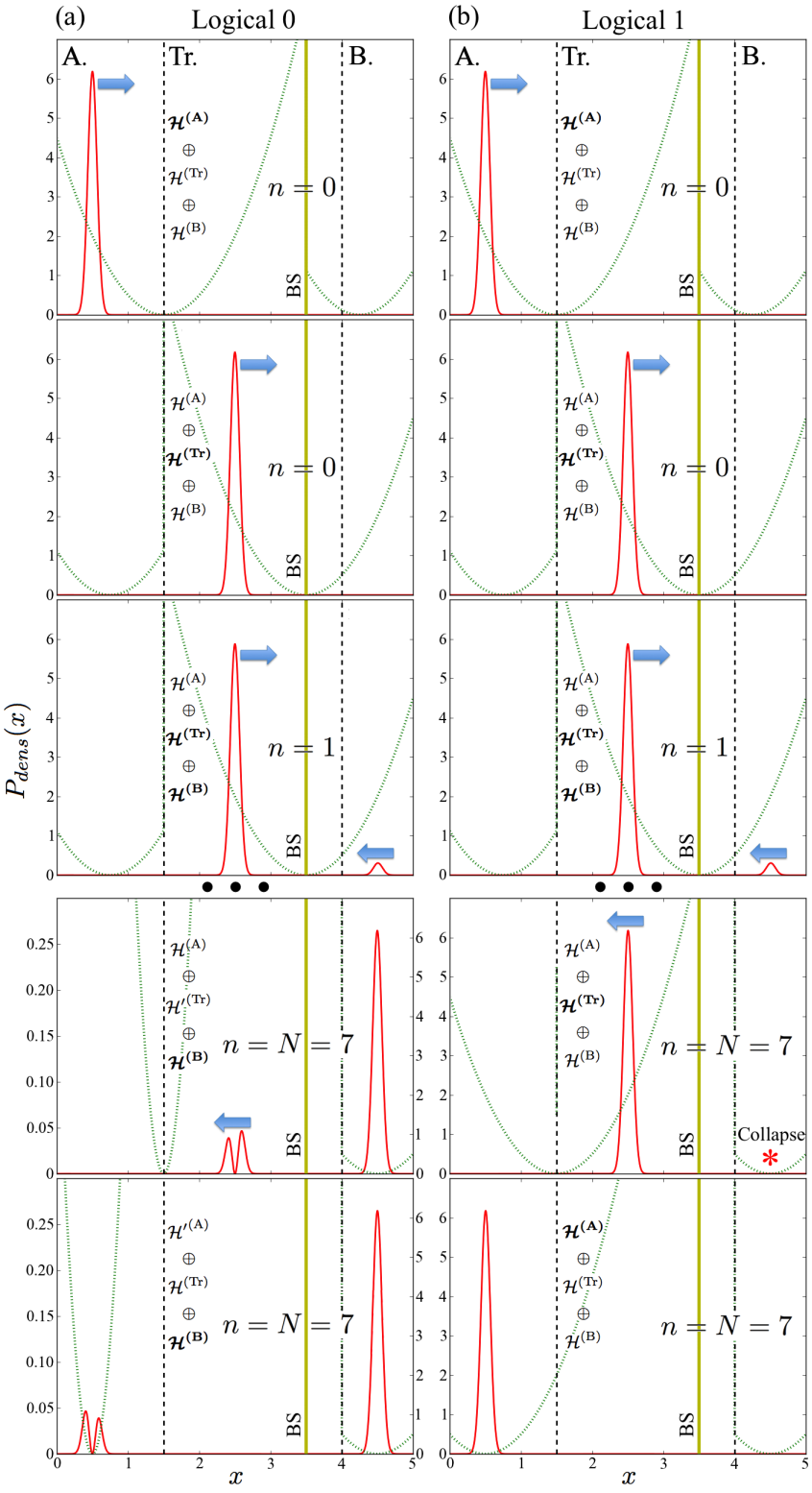


Figure 2.9 Type II CFC. The quantum evolution of the probability density distribution (solid red curve) as a function of x-position in the (a) 0-bit and (b) 1-bit processes. The $N = 7$ beam-splitter interactions take place at a potential barrier, indicated with a bronze yellow line at $x = 3.5$. The electrostatic potential is shown with dotted green lines. The spatial Hilbert spaces are separated by dashed black lines. The figure is further explained in the text. This figure originally appeared in Ref. [46].

generate a counterfactual violation in Alice's laboratory, as the wavefunction will annihilate at Bob's laboratory and cannot propagate back to Alice.

Finally, Fig. 2.9 makes it evident that unless a 0-bit error occurs, the wavefunction will never evolve from Bob's to Alice's laboratory. If we consider the 0-bit process with a weak measurement (or interaction) applied to the wavefunction in Bob's laboratory, this will encode some measurable impact in the particle wavefunction. However, if we assume high fidelity quantum channels and beam-splitters, the particle subject to the weak measurement will still end up in Bob's laboratory with a probability close to unity [see bottom frame of Fig. 2.9(a)]. Hence, the counterfactual nature of the protocol is stable against disturbances in Bob's laboratory. This is why the type II protocol has been described as counterfactual communication "without a weak trace" [46].

4.5 Logical bit-encoding and error correction

In Section 4.3, we saw that for a perfect quantum evolution, the probability of a successful 0-bit process will always equal unity, such that $P^0_{fail} = 0$, irrespective of the number of beam-splitters, N. The 1-bit process, however, fails with probability $P^1_{fail} = 1 - \cos^{2N}(\pi/2N)$; but such errors do not constitute counterfactual violations. This makes it possible to design an encoding scheme for logical bits, which keeps the protocol counterfactual at the same time as it brings the success rates of both the logical bit processes towards unity. Before doing this—and in the light of the numerical section above—let me make a small note with regard to realistic quantum evolutions.

Experimentally, it becomes increasingly difficult to keep the fidelity of the CMZI evolution high as N grows larger. For example, real beam-splitters will always have some error on the transmission and reflection coefficients, and there will always be some relative phase-changes between the different paths of an optical device. Hence, a finite error probability of the 0-bit process [visible in the two last frames of Fig. 2.9(a)] should be considered, such that $P^0_{fail} > 0$. Nevertheless, for reasonable values of N, and high-fidelity beam-splitters, the failure probability of the 0-bit process will be significantly smaller than that of the 1-bit process: $P^0_{fail} \ll P^1_{fail}$.

A useful bit-encoding scheme, or repetition code, can now be tailored using our knowledge about the error probabilities. Let each logical bit be decoded from a number, M, of processes (i.e. evaluations of the type II protocol). That is, Alice and Bob do not only need to have pre-determined time-slots at which Alice is to interrogate Bob via the protocol, but they also need to allocate M such slots per logical bit.

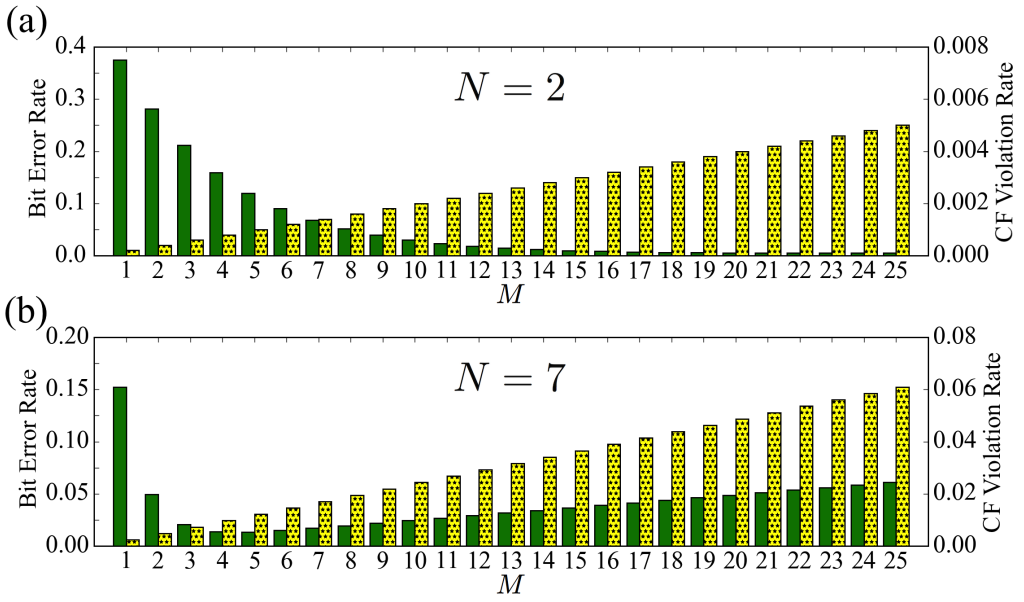


Figure 2.10 The green solid bars show the logical bit error rate, and the yellow dotted bars show the counterfactual violation rate, as functions of, M, the number of encoding processes. The beam-splitter angle, θ, has a mean of $\pi/2N$ rad and a standard deviation of $\sigma_\theta = 0.01$ rad. Figures (a) and (b) show simulations of devices with $N = 2$ and $N = 7$ beam-splitters, respectively. This figure originally appeared in Ref. [46].

Logical 0: Bob keeps his laboratory free of detectors for the M processes. If no detection is triggered in Alice's laboratory, she records a logical 0. The failure probability of this process is given by $P_{fail}^{\log 0} = 1 - (1 - P_{fail}^{0})^{M} \approx P_{fail}^{0} \times M$.

Logical 1: Bob keeps detectors in place for the M processes. If one or more detections are triggered in Alice's laboratory, she records a logical 1. The failure probability of this process is given by $P_{fail}^{\log 1} = (P_{fail}^{1})^{M}$.

Both logical failure events generate logical bit errors. However, owing to their respective wavefunction evolutions, only the failure of a 0-bit process constitutes a counterfactual violation. Consequently, our logical bit-encoding ensures that we can exponentially reduce the larger failure probability of the logical 1-bit process at the expense of a linear increase in the smaller failure probability of the logical 0-bit process.

Figure 2.10 shows how the number of bit processes, M, affect the bit error rate and the counterfactual violation rate for different numbers of beam-splitters ($N = 2, 7$). The bit error rate is taken to be the probability of a logical bit process giving the wrong value for a message with a balanced number of 0-bits and 1-bits. The counterfactual violation rate is then taken to be the probability (per logical bit process) of Alice receiving a particle which could have interacted with Bob's laboratory (see Chapter 3.8 for an in-depth discussion). The data has been produced by Monte Carlo simulations of the type II protocol with beam-splitters

of non-perfect θ values (standard deviation of 0.01 rad in Eq. 2.4). The data is built on simulations of 10^9 logical bit events (with a balanced number of 0s and 1s). For low values of N [Fig. 2.10(a)], the bit error rates are roughly exponentially decreasing with $(P^1_{fail})^M$ for $M < 18$. For larger values of N [Fig. 2.10(b)] the bit error rates quickly become linearly dominated by $P^0_{fail} \times M$ with increasing M.

It is clear that if high-fidelity quantum channels and beam-splitters are available, even small values of N enable successful reductions of the logical bit error rate in the protocol, whilst keeping the counterfactual violation rates to small fractions. Moreover, the bit-encoding scheme above can be extended to include other error sources as well as experimental losses. Work to experimentally realise this scheme is currently ongoing at the University of Vienna. The preliminary data supports the theory outlined in this chapter.

5 Ontological Implications

The first paper on the topic of IFMs suggested that they are a manifestation of quantum non-locality [59].[6] We have now seen that the concept of IFMs can be used in order to implement a single-particle protocol that can send one of two different bit-values counterfactually. This is in itself highly non-classical, as all previous communication protocols involve messages that propagate in the direction of the protocol particles. Let us now see if we can use the results of the type II protocol, and its peculiar arrangement of the spatial Hilbert spaces, to learn something more about the nature of quantum mechanics.

We start by considering the type II protocol from a more general perspective. The evolution of the quantum state in the 0-bit process can be described by a CMZI unitary transformation

$$|A\rangle \rightarrow \hat{U}\left(\hat{H}_{CMZI}\right)|A\rangle = |B\rangle\,, \tag{2.6}$$

where $|A\rangle \equiv |1\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|0\rangle^{(\mathrm{B})}$ and $|B\rangle \equiv |0\rangle^{(\mathrm{A})}|0\rangle^{(\mathrm{Tr})}|1\rangle^{(\mathrm{B})}$; and where the Hamiltonian that implements the CMZI evolution is represented by $\hat{H}_{CMZI}$. (Remember that the operations inside the individual spatial Hilbert spaces must be independent of one-another.) An example of such a Hamiltonian is the one considered in the toy-model simulation of Section 4.4. The Hamiltonian leads to the following propagation through the Hilbert space:

$$\mathscr{H}^{(\mathbf{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathbf{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \rightarrow$$

[6] Note that this form of quantum non-locality is different from the non-locality of the correlations of entangled particles discussed in, for example, Ref. [63].

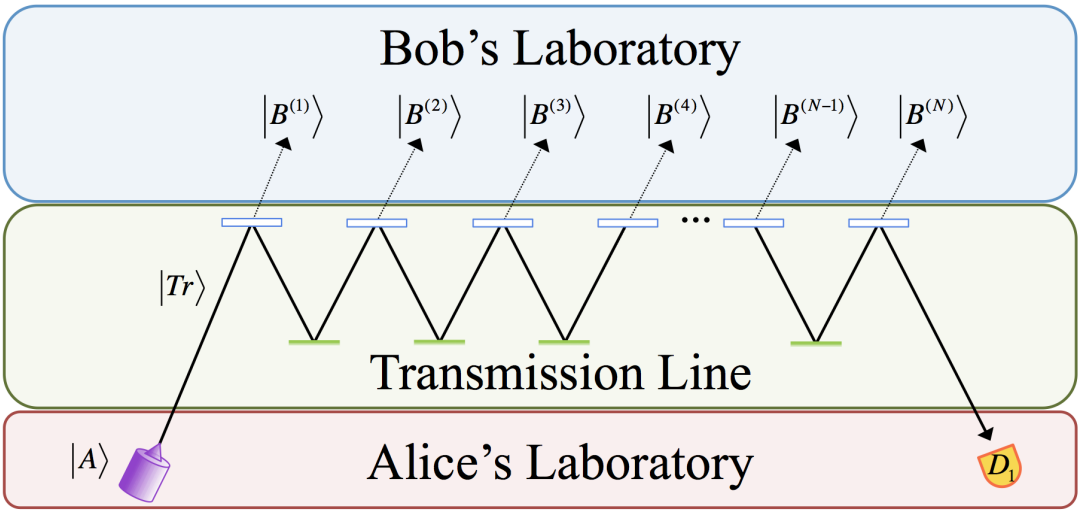


Figure 2.11 The 1-bit process of the type II protocol can be implemented by Bob removing the upper beam-splitters of the CMZI device, instead of inserting detectors as in Fig. 2.8.

$$\mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathbf{Tr})}} \oplus \boldsymbol{\mathscr{H}^{(\mathbf{B})}} \rightarrow \mathscr{H}^{(\mathrm{A})} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \boldsymbol{\mathscr{H}^{(\mathbf{B})}}.$$

In the Copenhagen interpretation, measurements are described by the non-unitary 4$^{\text{th}}$ postulate, which ensures that an observed particle is represented by a state corresponding to the measured eigenvalue. In the previous outline of the type II protocol (Section 4), the 1-bit process contains measurements, which make it difficult to compare it with the 0-bit process. However, as I showed in the final part of Chapter 1.6, it is possible to obtain the same effect as in the quantum Zeno effect, by extending the Hilbert space, without imposing any intermediate measurements. Instead of inserting detectors in his laboratory, Bob can remove the upper mirrors of the CMZI device. Figure 2.11 shows such a setup for the 1-bit process. The Hamiltonian of this scenario can be represented by: $\hat{H}_{CMZI} + \hat{H}_{(\mathrm{B})}$; where $\hat{H}_{(\mathrm{B})}$ is a Hamiltonian, local to Bob's laboratory, which removes the effect of the upper CMZI mirrors. In accordance with Fig. 2.11, the evolution of the 1-bit process described by a new unitary transformation:

$$\begin{aligned} |A\rangle \rightarrow & \hat{U}\left(\hat{H}_{CMZI} + \hat{H}_{(\mathrm{B})}\right)|A\rangle \\ & = \cos^N(\pi/2N)\,|A\rangle + \sum_{n=1}^{N} \cos^{n-1}(\pi/2N)\sin(\pi/2N)\,|B^{(n)}\rangle\,, \end{aligned} \tag{2.7}$$

where $|B^{(n)}\rangle$ represents the different single photon states that can occupy the n paths of Bob's laboratory (see Fig. 2.11).

Again, we take $N \to \infty$, such that the probability of occupying Bob's laboratory approaches zero. Figure 2.11 makes it clear that even if a part of the wavefunction were to enter Bob's laboratory, it could never return to the transmission line. The occupation of the spatial Hilbert spaces throughout the 1-bit process (see Section 4.3) is again given by

$$\begin{aligned}\boldsymbol{\mathscr{H}^{(\mathbf{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})} \to \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathbf{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \to \\ \mathscr{H}^{(\mathrm{A})} \oplus \boldsymbol{\mathscr{H}^{(\mathbf{Tr})}} \oplus \mathscr{H}^{(\mathrm{B})} \to \boldsymbol{\mathscr{H}^{(\mathbf{A})}} \oplus \mathscr{H}^{(\mathrm{Tr})} \oplus \mathscr{H}^{(\mathrm{B})}.\end{aligned}$$

Finally, we note that the only difference between the Hamiltonian applied in the 0-bit process of Eq. 2.6 and the 1-bit process of Eq. 2.7, is that the latter case includes the addition of Bob's local Hamiltonian, $\hat{H}_{(\mathrm{B})}$. Here, we are dealing with a single-particle system. In "standard" physics, a local Hamiltonian would affect a single particle via the interactions with force carriers only if the particle were within the domain of the Hamiltonian. The effect of the force carriers is determined by the specific Hamiltonian and the initial system. In this representation of the type II protocol, the effect of the local Hamiltonian, $\hat{H}_{(\mathrm{B})}$, is to preserve Alice's particle in the transmission line. But in the scenario, where Bob's local Hamiltonian is applied, the amount of wavefunction that occupies his laboratory is vanishingly small. Moreover, as the upper CMZI mirrors are removed, any wavefunction entering Bob's laboratory could not return to the transmission line. Consequently, the action of the local Hamiltonian causes an interaction-free non-local effect in terms of changing the evolutions of the protocol particle. This is irrespectively of the operations in Alice's laboratory and the transmission line being independent of Bob's action. The workings of the type II protocol are, perhaps, the most interesting manifestation of IFM non-locality. One could be tempted to interpret the results as the proof of the existence of a non-local "quantum force"—much like the one considered in the de Broglie-Bohm theory [14, 15].[7] The carriers of this force would be the individual components of the wavefunction, and the mediated "interaction" of the force would be the self-interference (or lack thereof) of them. As the need for such a force in the de Broglie-Bohm theory has been one of the arguments against it (see Ref. [20] for an overview), the fact that the Copenhagen interpretation also needs a similar "force" to account for counterfactual phenomena, adds new perspective to the debate about the ontology of quantum mechanics.

[7] In the de Broglie-Bohm theory the trajectory of a particle is deterministically determined by forces that arise from a local classical potential and a non-local quantum potential [14, 15].

6 Conclusion

There are two main proponents for the implementation of CFC. The first (type I) protocol, created by Salih *et al*., uses a complicated structure of many inner CMZIs nested within one outer CMZI. Given an infinite number of beam-splitters, and perfect fidelity of the quantum operations, the protocol seems to enable the transmission of information according to the type I definition of counterfactuality—that is, Bob can send Alice a message without the propagation of any particles between them. Nevertheless, the protocol is substantially flawed. Its requirement of perfect quantum channels is simply unrealistic; any experimental scenario would include small perturbations to the evolution of the wavefunction in the protocol. Moreover, there is a non-vanishing weak trace within Bob's laboratory, which, in the light of the TSVF discussion of Chapter 1, indicates counterfactual violations. Finally, the need for tens of thousands of perfectly tuned beam-splitters to implement the protocol with decent success probabilities, raises well-founded doubt about the possibility of experimentally realising the type I protocol.[8] In the following chapter, I shall provide a more thorough analysis of the type I protocol. The discussion will then be centred around an interpretation-independent analysis of the leakage of Fisher information from Bob's to Alice's laboratory.

The shortcomings of the first CFC protocol prompted me to develop a second (type II) protocol that uses a more flexible definition of counterfactuality, in which counterpropagation is allowed: Bob transmits information to Alice, in the opposite direction to the flow of particles.

The type II protocol utilises an IFM for only one of the bit-values. This is done with a CMZI device, without any nested MZI structures. This device simplification, together with a particular allocation of the individual spatial Hilbert spaces, proves adequate to overcome the problems faced by the type I protocol. Moreover, a specific logical bit-encoding scheme can be used to remove the need for a large number of optical components in the device. In fact, an MZI device with just two beam-splitters can be used to carry out the type II protocol with minute bit error and counterfactual violation rates. The protocol outlines how the mythological concept of telepathy can be realised with quantum mechanics.

Many optics-based protocols only consider the matrix evolutions of point-like particles. However, in order to draw conclusions about the counterfactual nature of a protocol, it is important to know the actual evolution of the spatial wavepackets involved. We have seen how such evolutions can be obtained via numerical staggered leapfrog simulations of the TDSE. Such simulations have previously been used to investigate Fermionic quantum information processing protocols. The same numerical framework proves itself useful in studying the

[8] $60,000$ beam-splitters (part number: TW1550R1A2) currently sell for roughly £$15,000,000$ at Thorlabs.

evolution of wavefunctions in optical interferometers. A Gaussian particle can be used in a toy-model to mimic wavefunction propagation in optical devices. The numerical study of the wavefunction evolution in the type II device shows that, for more realistic quantum channels and operations, there can be a slight "leakage" of the wavefunction from Bob to Alice. However, the impact of this is small on both the level of counterfactuality and the success-rate of the information transmission. I shall examine this further in the next chapter.

Whilst the type I protocol seems experimentally unattainable, the type II protocol is robust against experimental disturbances, both in terms of the ability to reduce bit errors and counterfactual violations. In conclusion, the type II protocol provides a rigid methodology which contradicts the intuitive idea and tenet of information theory, that a message must be carried by particles (or waves) that travel in the same direction as it does, and that the particles (or waves) that are received by the receiver must have interacted with the transmitter [21, 24]. This long-held premise of both classical and quantum physics should now be considered strictly classical. Finally, the type II protocol is a striking manifestation of interaction-free non-locality, which adds interesting new perspectives to the debate regarding the ontology of quantum mechanics.

Chapter 3

Particle Presence and Counterfactuality

1 Chapter Summary

In this chapter I shall outline a methodology for investigating how well a protocol satisfies type I or type II counterfactuality. The chapter is largely based on my paper from 2017, published in Physical Review A (Ref. [47]), and co-authored with Axel Gottfries and Crispin Barnes. In Section 3 I outline two information measures, the Shannon mutual information and the Fisher information, which can be used to investigate the suitability of a discrete quantum circuit for parameter estimation. In Section 4 I show how the classical Fisher information is related to the quantum evolution of the particle of interest. I shall argue that the Fisher information is precisely the measure that is needed to go beyond the Copenhagen interpretation and discuss inter-measurement particle presence. Section 5 contains a comprehensive study of parameter estimation with nested MZIs, the basic structures of type I counterfactual communication (CFC) devices. The Fisher information measure is then used to develop a methodology of how to evaluate type I counterfactual protocols in Section 6. This section also contains a methodology for the evaluation of type II protocols, based on the flow of probability densities. Finally, Section 7 and Section 8 contain thorough investigations of the counterfactual violations of the type I and type II protocols under the assumption of realistic, non-perfect, quantum channels.

2 Setting the Scene: Second Act

The study of interferometers has resulted in some of the most profound advances in physics. Their pivotal role is evident from a vast range of discoveries. An early example of this is the famous Michelson-Morley experiment, which established the speed of light as a constant [83].

A later, quantum example, is Hardy's paradox, which uses a simple interferometer device to demonstrate the non-local behaviour of quantum physics [64]. The same interferometer that was used in the Michelson-Morley experiment (the Michelson interferometer) is still used with great success today: the recent discovery of gravitational waves was made with two power-recycled Michelson interferometers [84].

As well as a way to explore and discover novel physics, interferometers are essential components in the detection of external fields and parameter estimation [32, 85–94]. A common example of this is phase estimation [30, 31, 89, 95]. In Chapter 1.3.4 we saw how a phase-shift can be detected with a Mach-Zehnder interferometer, whilst a free-space interaction with the phase-shifting medium is not detectable. Chapter 1.4 presented a simple example of how the maximum-likelihood function (Eq. 1.13) can be used to estimate phases. Often it is crucial that a large number of measurements be carried out for the phase estimation to be precise; the experiment has to be conducted several times to build up the output probabilities. A methodology for establishing how suitable a given interferometer experiment is for certain phase-estimation has been developed by Bahder *et al*. [30–32]. Their methodology is based on two measures from information theory: the Shannon mutual information, and the classical Fisher information, which are outlined in the Section 3.

As we have seen, both type I and type II definitions of CFC impose restrictions on the inter-measurement propagation paths of the protocol particles. This calls for a methodology for the analysis of inter-measurement particle presence.

In standard discussions of quantum mechanics, a physicist considering the Mach-Zehnder interferometer (see Fig. 1.5) is likely to regard the photon as existing in both the upper and lower interferometer arm between the two beam-splitters. However, as mentioned in Chapter 1, the Copenhagen interpretation does not provide us with any means of discussing such inter-measurement presence or physical reality. For more complicated structures than the MZI, there is yet no consensus on how this should be done, leading to intense debate.

In particular, criticism has been levelled at the claims of type I CFC protocols [36–38, 40–45]. In Chapter 2 we saw the essence of this criticism. It amounts to a discussion of the weak trace (see Chapter 1.5.3) which a quantum particle leaves within a nested Mach-Zehnder interferometer (NMZI) and how to use that to discuss inter-measurement presence.[1] However, the previous measures seem to hinge on specific interpretations of quantum mechanics, particular the two-state vector formalism (TSVF), briefly presented in Chapter 1.5.1. This chapter (as well as the next) is devoted to the development of a rigid

[1]The weak trace formalism is an interesting addition to the study of the foundations of quantum mechanics. Nonetheless, the physical meaning of the weak value, which is the basis of any calculation of a weak trace, has itself been subject to intense debate [45, 96, 97]. This has prompted the development of other investigations of inter-measurement quantum presence [98].

operational and interpretation-independent methodology, which enables the investigation of inter-measurement quantum presence, and ultimately, answers the question of whether or not a given interferometer protocol satisfies counterfactuality.

In order to do this, I shall have to depart from the Copenhagen interpretation. Instead, my methodology is based on the information theoretical principles that are rooted in the mathematics of parameter estimation of probabilistic processes. The underlying assumption of both counterfactuality measures presented in this chapter is that completely perfect quantum evolutions are impossible. More specifically, in reality there are no such things as absolutely lossless and pure quantum translations in space, and there are no perfectly exact unitary operations. As I stated in Ref. [47]: "a theory that relies on such perfect quantum channels is as valid (read: invalid) as a thermodynamic proof only valid at 0 K ". Instead of considering perfect quantum evolutions, we shall be considering evolutions under a simple noise model, mimicking the nature of physical disturbances. We shall see that these disturbances encode information in wavepackets propagating through interferometers. As this information is measurable at the output of an interferometer I shall argue for ways of using it to make inferences with respect to the physical past paths of the particles. In order to do this I shall extend the previous work on phase estimation, to incorporate general parameter estimation in more elaborate interferometers.

Neither of the interferometers we considered for CFC include polarization operations. However, any real photon propagation will include some alteration of the polarization degree of freedom. Hence, I suggest that a good model for disturbances in these interferometers can be based on minute polarization rotations. In particular, the classical Fisher information (i.e. the information obtained about the value of an unknown parameter—see Section 3.2) will turn out to be a useful tool in evaluating the disturbances, relating them to the particle presence in the interferometers of interest in this thesis. In Section 4 I shall outline a Fisher information measure which allows for the calculation of how much a particle "leaks" from the input of an interferometer to a specific part of it. (In Chapter 4 I extend this analysis and present a calculation of the quantum Fisher information from general quantum circuits subject to optimised measurements.)

As the type I protocol forbids the travel of Alice's particles to Bob's laboratory, the Fisher information measure can be used to evaluate the level of counterfactuality of this protocol. The type II protocol will also be evaluated, but with a counterfactuality measure tailored to its definition. The information theoretical study of this chapter suggests that we discard the type I CFC protocol. The type II protocol, on the other hand, will prove to be truly robust towards physical disturbances, and successfully satisfy its counterfactual definition even in the presence of noise.

3 Information Measures

One aim of this chapter is to provide two respective methodologies for evaluating whether or not a given protocol satisfies type I or type II counterfactuality.

The information measures I shall use allow us to see how a quantum wavefunction, propagating through an interferometer, transmits information from one spatial location to the output. In order to measure the information content of a particle at a given position, the unitary evolution of the wavefunction (and thus also the probability density) in the device of interest must be known. As in the previous chapters, we represent the evolution of our discrete photon state, $|\psi_{in}\rangle$, by the operation of a unitary matrix: $|\psi_{in}\rangle \rightarrow |\psi_{out}\rangle = \hat{U}\,|\psi_{in}\rangle$. Through this evolution other particles or external media will interact with the wavefunction; effectively, these interactions encode information in the probability density distribution of the quantum state.

The Copenhagen interpretation makes no claims about the quantum wavefunction being an element of *physical reality*, but merely states that it is an abstract function which allows us to calculate the outcome of *real* experiments with striking precision.[2] It is often considered nonsensical to ask where a quantum particle was present between the two observations corresponding to the pre-selected initial state and a post-selected state obtained in a final measurement; the basic idea being that inter-measurement observations also constitute measurements and thus changes the quantum state of interest via the 4^{th} postulate. However, in Chapter 1.5.3 we saw attempts of how to address this question in the light of weak measurements. In this chapter I shall continue to argue that there are ways to defy the founding fathers of quantum mechanics, and interrogate quantum states about their previous whereabouts. Furthermore, the ability to discuss inter-measurement presence is required if we are to evaluate counterfactuality, which hinges on restrictions of inter-measurement trajectories. Hence, if we want to be able to draw conclusions about the impact of CFC on quantum theory, we must have a measure for inter-measurement presence. If such a measure can be designed in an interpretation-independent way, the conclusions would naturally be stronger.

The basic component behind my presence measure is the inevitable occurrence of wavefunction disturbances and unwanted weak interactions. Consider a quantum particle, evolving from input to output. Along this evolution let there be an interaction with an external medium which causes the wavefunction to occupy parts of the Hilbert space that are only available via this interaction. The wavefunction will subsequently carry some *information* about the nature of this interaction. It is therefore possible to argue that the components of the probability

[2] Prior to the development of quantum mechanics, discussions were based on the classical tenet that a physical theory should correspond to an underlying physical reality.

density that are output in a state—only available via the interaction—have had a previous *presence* at the spatial location of the interacting medium.

In order to use the idea above to build a measure for inter-measurement presence, we need a way of quantifying the information content that is encoded in a wavefunction. In Chapter 1.4 we saw how knowledge of the output probability distribution can be used to estimate the interaction parameter of a phase-shifter. The following sections outline quantitative measures of the effectiveness of such parameter estimation, providing ways of quantifying the information gain in interferometric experiments. This theory is applicable to any multi-photon state. However, in order to keep in line with the CFC protocols [39, 42, 46, 78], we shall limit this chapter to single-photon scenarios.

3.1 Shannon mutual information

Let us consider an interferometer which is known in detail, except for a single (continuous) parameter, θ, which sets some interaction along the device. This interaction can be the result of a component of the interferometer, or of an interaction with an external medium. An experimentalist inserts the input state, ψ_{in}, and subsequently makes a measurement at the output paths of the interferometer. The suitability of this experiment for the estimation of the parameter θ is given by the Shannon mutual information, $H(\theta : \mathscr{M})$, between the parameter θ and the $\mathscr{M} = \{\mathscr{M}_i\}$ (discrete) set of measurement outcomes of the experiment. $\mathscr{M}_i$ represents a measurement outcome that occurs at the i^{th} detector of the total spatial Hilbert space, $\mathscr{H} = \{\mathscr{H}_i\}$. The Shannon mutual information is given by

$$H(\theta : \mathscr{M}) = \sum_i \int_{\theta_{min}}^{\theta_{max}} d\theta P(\mathscr{M}_i|\theta, \psi_{in}) p(\theta) \times \log_2 \left[\frac{P(\mathscr{M}_i|\theta, \psi_{in})}{\overline{P}(\mathscr{M}_i|\psi_{in})} \right], \tag{3.1}$$

where $\overline{P}(\mathscr{M}_i|\psi_{in}) = \int_{\theta_{min}}^{\theta_{max}} d\theta' P(\mathscr{M}_i|\theta', \psi_{in}) p(\theta')$; $P(\mathscr{M}_i|\theta, \psi_{in})$ is the probability of $\mathscr{M}_i$ for some specific θ and ψ_{in}; and $p(\theta)$ is the *a priori* probability distribution of the parameter θ, that is, $p(\theta)$ reflects the experimentalist's knowledge of θ prior to the experiment [30, 99]. The Shannon mutual information thus gives a measure of how much information of θ the experimentalist can obtain via the knowledge of the measurement outcomes of repeated evaluations of the experiment. The higher the value of $H(\theta : \mathscr{M})$ the better the experiment is for the parameter estimation of θ.

3.2 Classical Fisher information

Let us now move on to the study of Fisher information. Consider the same type of experiment as above, with measurement outcomes $\mathscr{M} = \{\mathscr{M}_i\}$. We want to calculate how well the experiment allows us to estimate the unknown parameter θ given that it has a fixed definite value. The classical Fisher information is given by

$$F(\theta) = \sum_i \frac{1}{P(\mathscr{M}_i|\theta,\psi_{in})}\left[\frac{\partial}{\partial\theta}P(\mathscr{M}_i|\theta,\psi_{in})\right]^2. \tag{3.2}$$

The quantity $F(\theta)$ is a measure of the average quantity of information of the specific value of the unknown parameter, θ, which the experimenter obtains per evaluation of the experiment [32, 99]. The Fisher information thus grows linearly with the number of repetitions of the experiment. In parameter estimation (see Chapter 1.4), the Fisher information is closely related to the variance of the unbiased estimator, θ_e, via the Cramér-Rao bound:

$$\mathrm{Var}(\theta_e) \geq \frac{1}{F(\theta)}. \tag{3.3}$$

A simple and beautiful derivation of this relation can be found in Ref. [100].

A short note on the differences between Shannon mutual information and classical Fisher information: the former measures how good an experiment is for the estimation of the unknown parameter θ, whilst the latter gives a measure of how much information of a specific θ can be obtained through the experiment.

3.3 Quantum Fisher information

The classical Fisher information, Eq. 3.2, is calculated with respect to some probability output distribution as produced by a specific measurement conducted at the end of an experiment. This is the case for interferometers with well-defined detector positions, investigated in this and previous chapters. The quantum Fisher information, $F_Q(\theta)$, however, considers an output quantum state, ρ_θ (now a density operator), prior to the measurement of it. $F_Q(\theta)$ gives us the maximum classical Fisher information we can obtain of θ, given that we measure ρ_θ with the most suitable generalised measurement (positive-operator valued measure) [101–104]. The quantum Fisher information is defined as follows:

$$F_Q(\theta) = \mathrm{Tr}[\rho_\theta \Lambda_\theta^2], \tag{3.4}$$

where Λ_θ is the logarithmic derivative operator, implicitly defined by

$$\frac{\partial}{\partial\theta}\rho_\theta = \frac{1}{2}(\Lambda_\theta\rho_\theta + \rho_\theta\Lambda_\theta). \tag{3.5}$$

The quantum Fisher information thus places an upper bound on the classical Fisher information, and a further lower bound on the variance of the estimator, θ_e:

$$\mathrm{Var}\big(\theta_e\big) \geq \frac{1}{F(\theta)} \geq \frac{1}{F_Q(\theta)}. \tag{3.6}$$

4 Fisher Information and Inter-Measurement Presence

A large number of quantum interferometers used in experiments or for theoretical discussions do not contain any polarization rotations. None of the interferometers discussed in the previous chapters have included polarization manipulations. Nevertheless, any "real" experiment with photons will contain some, however minute, polarization rotations. They might be the result of unwanted interactions, impurities in the interferometer materials, or systematic errors in the experimental setup. In this chapter, the imperfect nature of *real* quantum channels is mimicked by the introduction of a single unwanted polarization rotation, placed somewhere in the optical circuit.

Whilst we consider the interferometer to be assigned to the laboratory of a primary experimenter, Alice, the location in the vicinity of the polarization rotation is a part of the laboratory of a secondary external experimenter, Bob. The polarization rotation thus represents a disturbing interaction with Bob's laboratory, rather than a generic noise model.[3] We refer to this interaction as the "tagging" of the wavefunction.

To accommodate for this polarization tagging we must extend the Hilbert space to include a polarization degree of freedom. The Hilbert space of our system is thus formed from a spatial component and a polarization component: $\mathscr{H} = \mathscr{H}_S \otimes \mathscr{H}_P$. The classical Fisher information can then be used to estimate the parameters associated with the tagging interaction. This tells us how much information of the interaction in Bob's laboratory is available to Alice at the output of her interferometer.

In the previous section I suggested that a tagged part of a wavefunction can be considered to have been previously present at the location of the tagging interaction. In this section I shall justify this further. We shall see that if all the output states of the interferometers studied in this thesis are accessible to Alice, the classical Fisher information she obtains with respect

[3] The unwanted weak polarization rotations can occur at any part of a realistic device. However, we limit our interest to a small spatial domain controlled by Bob (central rectangle in Fig. 3.1), and examine only single unwanted polarization rotations. Chapter 4 contains a more general analysis.

to Bob's interaction, is directly proportional to the integrated probability density distribution that has propagated through his laboratory in the Schrödinger picture.

First, we introduce an optical input vector, $\boldsymbol{a}$, which evolves to an output vector, $\boldsymbol{b}$. Both vectors are taken to be of length $2n$, corresponding to n different spatial levels of the interferometer, each of which can exist in one of the two polarization states. Moreover, we order the vector elements such that the first n correspond to states with the same polarization as the input state (taken to be uniform over potentially spatially separated input states). The following n elements have orthogonal polarization to the input state.

Now, the evolution of the input vector, $\boldsymbol{a}$, to the output vector, $\boldsymbol{b}$, can be described by a scattering matrix, $\mathscr{S}$, composed of three unitary operations:

$$\mathscr{S}\boldsymbol{a} \equiv \left[\hat{V}\cdot\hat{f}^{(k)}(\theta)\cdot\hat{U}\right]\boldsymbol{a} \equiv \boldsymbol{b}^{(k)}, \tag{3.7}$$

where the unitary evolutions up to and after the tagging interaction are described by $\hat{U}$ and $\hat{V}$, respectively; and where the unitary evolution of the *single* polarization tagging interaction, set by the parameter θ at the spatial location of k ($1 \leq k \leq n$), is given by $\hat{f}^{(k)}(\theta)$. Note that $\hat{H}$, $\hat{V} \in u(\mathscr{H}_S)$ and $\hat{f}^{(k)}(\theta) \in u(\mathscr{H}_P)$. Figure 3.1 shows a sketch of the optical circuit of Eq. 3.7. A single-photon state with n spatial components, of uniform polarization (defining the horizontal axis), is input to the optical circuit. The quantum state evolves through the optical circuit, after which single photon polarization measurements of the n spatial output ports are made. As can be seen from the figure, the unitary operations $\hat{U}$ and $\hat{V}$ do not act on the polarization degree of freedom, in accordance with the interferometers studied in this thesis.

In order to obtain a relationship between the classical Fisher information and the integrated probability distribution at the location of the interaction, we need to work through a number of steps.

1. The initial wavefunction, $\boldsymbol{a}$, is first acted upon by the unitary operation, $\hat{U}$. The wavefunction after this operation can be defined as $\boldsymbol{c} \equiv \hat{U}\boldsymbol{a}$, with

$$c_i = \sum_{j=1}^{2n} U_{i,j}a_j. \tag{3.8}$$

2. After this, the wavefunction is subject to the tagging rotation. This operation rotates the polarization state of the wavefunction between the vector elements k and k', and

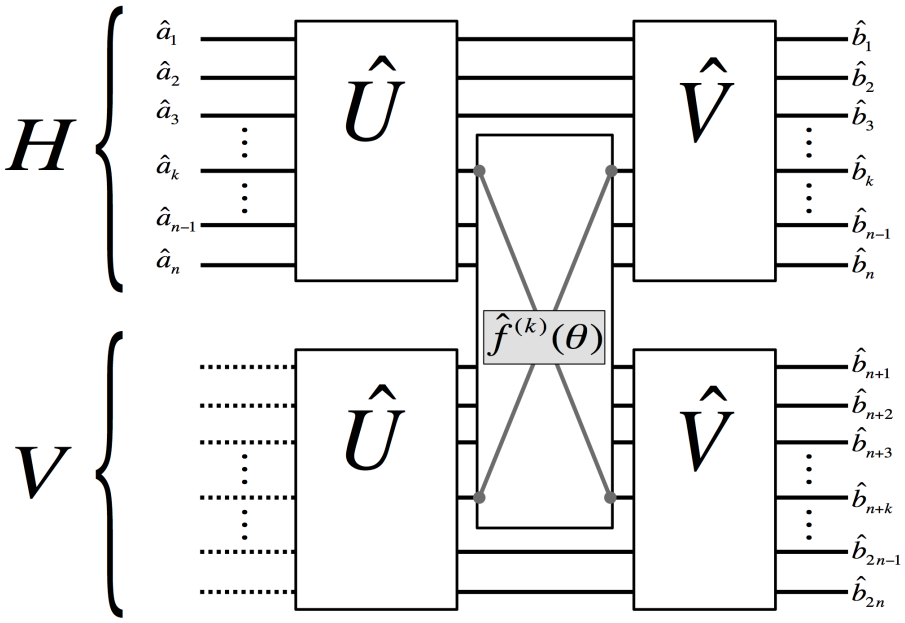


Figure 3.1 Sketch of an optical circuit of the form of Eq. 3.7. A single-photon state with horizontal polarization is initialised in a superposition of spatially separated paths ($\hat{a}_i$ with $i \in \{1,...,n\}$). The spatial part of the wavefunction is then subject a unitary operation, $\hat{U}$. After $\hat{U}$ the wavefunction component at a single path is subject to a tagging polarization rotation, $\hat{f}^{(k)}(\theta)$, which brings it into a superposition of horizontal and vertical polarization. Finally, the spatial part of the wavefunction is subject to a unitary operation, $\hat{V}$, after which each spatial output path is measured in the linear polarization basis.

can be represented by a real unitary matrix:

$$\hat{f}^{(k)}(\theta) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & f_{k,k}(\theta) & \cdots & f_{k,k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & f_{k',k}(\theta) & \cdots & f_{k',k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \tag{3.9}$$

where $f_{k,k}(\theta) = \sqrt{1 - \left[f_{k,k'}(\theta)\right]^2}$, $f_{k',k}(\theta) = -f_{k,k'}(\theta)$ and $f_{k',k'}(\theta) = f_{k,k}(\theta)$, with $f_{k,k}(\theta) : \mathbb{R} \to [-1,1]$.

3. Without loss of generality we can require that the order of the vector elements is such that element l and $l+n$ correspond to the same spatial location, for $1 \leq l \leq n$. This effectively ensures that $c_i = 0$ for $i > n$ in Eq. 3.8, and that $k' = k+n$ in Eq. 3.9.

4. Using the imposed vector element ordering we express the quantum state after the tagging interaction as $\boldsymbol{d}^{(k)}(\theta) \equiv \hat{f}^{(k)}(\theta)\boldsymbol{c}$, with

$$d_i^{(k)}(\theta) = \sum_{j=1}^{2n} f_{i,j}^{(k)}(\theta) c_j. \tag{3.10}$$

   As a consequence of the vector element ordering, the only components of $\boldsymbol{d}^{(k)}$ that depend on θ are $d_{i=k}^{(k)} = f_{k,k}(\theta)c_k$ and $d_{i=k+n}^{(k)} = f_{k+n,k}(\theta)c_k$.

5. Following the rotation matrix, the final unitary operation, $\hat{V}$, is applied to the quantum state. In accordance with the steps above, we express the circuit output state as $\boldsymbol{b}^{(k)}(\theta) \equiv \hat{V}\boldsymbol{d}^{(k)}(\theta)$, with

$$b_i^{(k)}(\theta) \equiv \beta_i^{(k)} + b_{i,\theta}^{(k)}(\theta) = \sum_{j=1}^{2n} V_{i,j} d_j^{(k)}(\theta), \tag{3.11}$$

   where the θ-dependence of $b_i^{(k)}(\theta)$ is captured by $b_{i,\theta}^{(k)}(\theta)$, and where a corresponding θ-independent term, is defined as $\beta_i^{(k)}$.

6. We can now express the probability of detecting the single-photon in the i^{th} output:

$$P_i(\theta) = \left|\beta_i + b_{i,\theta}(\theta)\right|^2, \tag{3.12}$$

   where we drop the (k)-superscript, for convenience. This probability can be re-expressed as

$$P_i^{(k)}(\theta) \equiv \left|\beta_i^{(k)} + b_i^{\prime(k)} f_{\chi,k}^{(k)}(\theta)\right|^2, \tag{3.13}$$

   where $\chi = k$ if $i \leq n$ and $\chi = k+n$ if $i > n$.

7. Using Eq. 3.2, the individual classical Fisher information components are given by

$$\begin{aligned} F_i =& \frac{1}{\left|\beta_i + b_i' f_{j,k}(\theta)\right|^2} \left[\frac{\partial}{\partial\theta}\left|\beta_i + b_i' f_{j,k}(\theta)\right|^2\right]^2 \\ =& \frac{1}{\left[\beta_i + b_i' f_{j,k}(\theta)\right]\left[\beta_i^* + b_i'^* f_{j,k}(\theta)\right]} \\ & \times \left\{\frac{\partial}{\partial\theta}\left[\beta_i + b_i' f_{j,k}(\theta)\right]\left[\beta_i^* + b_i'^* f_{j,k}(\theta)\right]\right\}^2. \end{aligned} \tag{3.14}$$

We can simplify this expression by expressing the coefficients as $\beta_i \equiv |\beta_i| e^{i\phi_i}$ and $b_i' \equiv |b_i'| e^{i\phi_{i,\theta}}$ and defining $\Phi_i^I \equiv \phi_i - \phi_{i,\theta}$:

$$\begin{aligned} F_i =& \frac{\left[\cos(\Phi_i^I)|\beta_i| + |b_i'| f_{j,k}(\theta)\right]^2}{|\beta_i|^2 + |b_i'|^2 f_{j,k}^2(\theta) + 2\cos(\Phi_i^I)|\beta_i||b_i'| f_{j,k}(\theta)} \times 4|b_i'|^2 \left[\frac{\partial}{\partial\theta} f_{j,k}(\theta)\right]^2 \\ \equiv& K_{i,\theta}(\theta, \Phi_{i,\theta}^I) \times 4|b_i'|^2 \left[\frac{\partial}{\partial\theta} f_{j,k}(\theta)\right]^2, \end{aligned} \tag{3.15}$$

where $K_{i,\theta}(\theta, \Phi_{i,\theta}^I)$ is a phase-dependent coefficient ranging between 0 and 1.

8. The expression of $K_{i,\theta}(\theta, \Phi_{i,\theta}^I)$ is rather convoluted. Nevertheless, we note that $K_{i,\theta}(\theta, \Phi_{i,\theta}^I = m\pi) = 1$, for $m \in \mathbb{Z}$. Actually, a vast number of interesting interferometers satisfy this phase criterion. It is, for example, always satisfied if the spatially separated wavefunction components that are input to the interferometer of Fig. 3.1 share the same phase, and $\mathscr{S}$ can be described by a real matrix. Moreover, $K_{i,\theta}(\theta, \Phi_{i,\theta}^I) = 1$ for the optical setups considered previously in this thesis and for those discussed in Refs. [36, 39, 46, 47, 59, 75, 76]. The individual Fisher components of the optical setups considered so far in this thesis can, consequently, be simplified such that

$$F_i = 4|b_i'|^2 \left[\frac{\partial}{\partial\theta} f_{j,k}(\theta)\right]^2. \tag{3.16}$$

Using Eq. 3.2 we find that the total Fisher information is given by

$$F(\theta) = \sum_{i=1}^{2n} 4\left[\frac{\partial}{\partial\theta}\left|b_{i,\theta}(\theta)\right|\right]^2. \tag{3.17}$$

9. We can expand Eq. 3.17 in terms of the unitary operation elements of $\hat{V}$:

$$\begin{aligned} F(\theta) =& \sum_{i=1}^{n} 4\left[\frac{\partial}{\partial\theta}\left|V_{i,k} d_k(\theta)\right|\right]^2 + \sum_{i=n+1}^{2n} 4\left[\frac{\partial}{\partial\theta}\left|V_{i,k+n} d_{k+n}(\theta)\right|\right]^2 \\ =& \sum_{i=1}^{n} 4\left|V_{i,k}\right|^2\left[\frac{\partial}{\partial\theta}\left|d_k(\theta)\right|\right]^2 + \sum_{i=n+1}^{2n} 4\left|V_{i,k+n}\right|^2\left[\frac{\partial}{\partial\theta}\left|d_{k+n}(\theta)\right|\right]^2. \end{aligned} \tag{3.18}$$

$\hat{U}$ and $\hat{V}$ do not operate on the polarization degree of freedom of the quantum wavefunction. Hence, the symmetry of $\hat{V}$ is such that $V_{i,k} = V_{i+n,k+n}$. The symmetry of $\hat{V}$ also implies that $V_{i,k} = 0$ for $i > n$ and $V_{i,k+n} = 0$ for $i \leq n$. Remembering that we have assumed a real $\mathscr{S}$ we define a suitable reference-point for the uniform global phase

of the input quantum state. The expression for the classical Fisher information then simplifies further:

$$F(\theta) = \sum_{i=1}^{2n} 4|V_{i,k}|^2 \Big[\frac{\partial}{\partial\theta}|d_k(\theta)|\Big]^2 + \sum_{i=1}^{2n} 4|V_{i,k}|^2 \Big[\frac{\partial}{\partial\theta}|d_{k+n}(\theta)|\Big]^2. \tag{3.19}$$

10. The square of the absolute value of all the row elements of a unitary matrix sums to unity. The expression again simplifies:

$$F(\theta) = 4\Big[\frac{\partial}{\partial\theta}|d_k(\theta)|\Big]^2 + 4\Big[\frac{\partial}{\partial\theta}|d_{k+n}(\theta)|\Big]^2. \tag{3.20}$$

At this point we can see that the total Fisher information is independent of the unitary operation, $\hat{V}$, which follows the tagging interaction.

11. Finally, we substitute back the expressions of d_k and d_{k+n} from Eq. 3.10. This gives us the following equation for the classical Fisher information:

$$\begin{aligned} F^{(k)}(\theta) =& 4|c_k|^2 \left(\Big\{\frac{\partial}{\partial\theta}\big[f_{k,k}(\theta)\big]\Big\}^2 + \Big\{\frac{\partial}{\partial\theta}\big[f_{k+n,k}(\theta)\big]\Big\}^2 \right) \\ =& 4|c_k|^2 \frac{\Big[\frac{\partial}{\partial\theta} f_{k+n,k}(\theta)\Big]^2}{\big[f_{k,k}(\theta)\big]^2}, \end{aligned} \tag{3.21}$$

where the (k)-superscript has been re-introduced.

$F(\theta)$ in Eq. 3.21 satisfies the equality relation to the quantum Fisher information (right hand side of Eq. 3.6). This is highlighted in Chapter 4, where the above analysis is extended to a general quantum circuit. Hence, from an information acquiring perspective, the polarization measurements at each individual output path are optimal for the devices we are interested in.

Discussion

First, we note that $|c_k|^2$ in Eq. 3.21 represents the probability of observing the single-photon state at the k^{th} spatial level if a detector had been inserted where the tagging interaction takes place. Or equally, it is the integrated probability density distribution at this location. Equation 3.21 shows that the classical Fisher information (of the interferometers of interest) is directly proportional to this value: $|c_k|^2$. A naïve argument could say that one can use the extent that the wavefunction spreads into a certain spatial location, according to the time-dependent Schrödinger equation, as a measure of the extent that the quantum particle has had a *presence*

there. However, given the abstract nature of the complex wavefunction and the peculiarities of many quantum phenomena, it is philosophically difficult to specify how the *reality* of the wavefunction should be interpreted between measurements.[4]

Nevertheless, in the interferometers we are interested in investigating, the tagging interaction is the only operation that alters the polarization state of the wavefunction, rotating it from its initial state of polarization into a superposition state. It is arguably less problematic to consider an output photon, found in a polarization state different from the initial one, to have had a past that has included a spatial propagation through the location of the tagging mechanism of the circuit. Equation 3.21 shows that the information content, propagating from the tagging interaction to the output ports of the quantum circuit, is weighted by the integrated probability density distribution at the location of the tagging. This is true even if the tagging polarization rotations, which we have introduced to mimic unwanted disturbances, are of vanishing strengths. Owing to this direct proportionality, and the fact that the classical Fisher information represents the *real* and *measurable* information gain at the output of the circuit, I argue that the Fisher information (Eq. 3.21) is a good measure of the *extent* to which a particle has had a presence at the location of the tagging interactions in interferometers that satisfy the assumptions made above.

If an interferometer satisfies all other assumptions made in this section, but includes operations on the polarization degree of freedom (see, for example, the Michelson-based device in Ref. [39]), these operations should be included in $\hat{U}$ and $\hat{V}$. It is then more appropriate to calculate the Fisher information with respect to tagging mechanisms on an alternative degree of freedom. The crucial point is that this degree of freedom is unaffected by the ideal interferometer. With adequate alterations of the model above, the tagging operation, $\hat{f}^{(k)}(\theta)$ can, for example, be assigned to a rotation of the internal angular momentum of the photons. The beauty of this Fisher information analysis is that Eq. 3.21 will still be valid.

Let me highlight some fundamental differences between this methodology for evaluating inter-measurement presence and the one based on the weak trace, which we studied in Chapter 1.5.3 and Chapter 2.3.2. The interpretation of the weak trace as a measure of particle presence relies on the assumption that the weak value measures a real physical property of the inter-measurement quantum particle. Whether or not this is the case has been debated intensively (see Refs. [97, 106–108] for examples). The Fisher information measure, on the other hand, simply quantifies the flow of information from the tagging part of a quantum circuit to the output. As the information is encoded in the quantum particle it is arguably less controversial to relate this information to the inter-measurement presence of the particle. This

[4]For a lengthy discussion on this, see Wheeler's "The 'Past' and the 'Delayed-Choice' Double-Slit Experiment", which has been reproduced in Ref. [105].

is discussed in detail in Chapter 4. Another difference is that the weak trace measure requires the consideration of a post-selected quantum state that travels backwards in time, whilst the Fisher information measure is operational and simply concerns itself with the observed probabilities of the measurement outcomes. These outcome probabilities can be post-selected and renormalised (again see Chapter 4 for more details), but at no point does this require the consideration of a backwards propagating quantum state. Finally, the calculation of a weak value requires a disturbing weak interaction of finite strength, whilst the Fisher information can be calculated even when the interaction strength is zero.

Appendix B contains a more detailed analysis of the classical Fisher information in general quantum optical circuits, without assumptions regarding the phase of the input wavefunction, what degrees of freedom to include in the circuit, or restrictions on which quantum levels the tagging interaction acts on. However, the CFC devices investigated in this thesis are all non-polarizing in their ideal limit, and for now we proceed with calculations of the Fisher information with respect to polarizing tagging mechanisms.

5 Information in NMZIs

The basic structure of the type I CFC protocol is the NMZI device. The evolution of quantum particles in this type of structure has been studied and debated intensely over the last few years (for examples see Refs. [36, 45]). Whilst parameter estimation with MZIs has been scrutinised significantly [30, 32], a rigorous study of the information encoding capacity of NMZIs has been lacking from the literature. Before we proceed to the study of counterfactuality, I present my findings of NMZI parameter estimation, recently published in Ref. [47].

5.1 Free-space interaction

In the section above, we saw that the Fisher information of an interaction is proportional to the square of the wavefunction component that passed through it. Hence, a suitable reference scenario for a general circuit is that of a free-space interaction, where the entire wavepacket travels through the interaction. Our simple reference scenario considers a single-photon state, propagating in free space through a polarization rotator. Shortly after the interaction with the polarizing medium (the rotator) the quantum state is measured. This is depicted in Fig. 3.2.

The interaction of the quantum particle with the polarization rotator is set by the parameter θ. Using the measurement outcomes it is possible to calculate the classical Fisher information, $F(\theta)$, and the Shannon mutual information, $H(\theta : \mathscr{M})$.

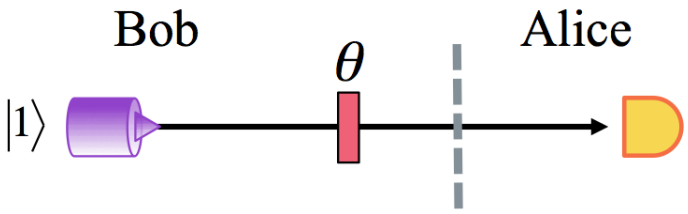


Figure 3.2 A single-photon state propagates through a polarization rotator after which it is subject to number state and polarization state measurements. The grey line separates Alice's and Bob's respective laboratories.

Let us consider the bosonic creation operators, $\hat{a}_H^\dagger$ and $\hat{a}_V^\dagger$, which create photons in horizontal and vertical polarization states, respectively. Moreover, we can define the polarization axes such that the single-photon input can be expressed as

$$\hat{a}_H^\dagger |0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{3.22}$$

In the reference scenario of Fig. 3.2 we have that: $\hat{U} = \hat{V} = \hat{1}$, and the scattering matrix (Eq. 3.7) simply takes the form of the tagging interaction, which we define to be given by

$$\mathscr{S}_{free} = \hat{f}^{(1)}(\theta) = \begin{pmatrix} \sqrt{1-\theta^2} & \theta \\ -\theta & \sqrt{1-\theta^2} \end{pmatrix}. \tag{3.23}$$

Throughout this chapter, we shall continue to model tagging interactions by applying unitary operations of this form.

The detector shown in Fig. 3.2 conducts measurements in the $|n_H, n_V\rangle$ basis, where n_H and n_V are the respective numbers of photon found in the horizontal and vertical state at the location of the detector.

The output probabilities of the experiment are given by

$$P(n_H = 1|\theta) = 1 - \theta^2, \tag{3.24}$$
$$P(n_V = 1|\theta) = \theta^2, \tag{3.25}$$

where the notation is such that $n_x = 1$ implicitly ensures that all other possible measurement outcomes, $y \neq x$, satisfy $\sum_{y\neq x} n_y = 0$ (as we only consider single-photon input states). I shall continue to use this notation throughout this thesis.

We are now in position to calculate the Shannon mutual information and the classical Fisher information with respect to the unknown parameter θ. Let us start with the former. If we assume no prior knowledge of θ, such that $\theta_{min} = -1$, $\theta_{max} = 1$ and $p(\theta) = 1/2$ in Eq.

3.1, the mutual information is given by

$$H_{free}(\theta : \mathscr{M}) = \frac{\ln(108) - 4}{3\ln(2)} \approx 0.328. \tag{3.26}$$

If we instead consider a fixed θ, the classical Fisher information of the free-space rotation can be calculated by either Eq. 3.2 or Eq. 3.21:

$$F_{free} = \frac{4}{1-\theta^2}. \tag{3.27}$$

Note that the Fisher information diverges for $\theta \to \pm 1$. This is because at $\theta = \pm 1$ the information gained from the $n_H = 1$ outcome tends to infinity at a quadratically faster rate than $P(n_H = 1|\theta)$ tends to zero. (At $\theta = 0$ the information gained from the $n_V = 1$ outcome also tends to infinity, but $P(n_V = 1|\theta)$ tends to zero at the same rate.) Further on in this chapter we shall use Eq. 3.27 as a free-space benchmark Fisher information when calculating counterfactual violation strengths.

5.2 Nested MZI interaction

Consider the NMZI device shown in Fig. 3.3. It contains four beam-splitters described by Eq. 1.2:

$$U_i = \begin{pmatrix} r_i & t_i \\ -t_i & r_i \end{pmatrix}, \tag{3.28}$$

where i labels the beam-splitter.

The normalised input and output vector of an NMZI can, in general, be described by $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively:

$$\boldsymbol{a} = \frac{1}{\sqrt{\mathscr{N}_a}} \begin{pmatrix} a_{1,H} \\ a_{2,H} \\ a_{3,H} \\ a_{1,V} \\ a_{2,V} \\ a_{3,V} \end{pmatrix}, \boldsymbol{b} = \frac{1}{\sqrt{\mathscr{N}_b}} \begin{pmatrix} b_{1,H} \\ b_{2,H} \\ b_{3,H} \\ b_{1,V} \\ b_{2,V} \\ b_{3,V} \end{pmatrix}, \tag{3.29}$$

where $\mathscr{N}_a$ and $\mathscr{N}_b$ are some normalisation constants. The i^{th} quantum level has a corresponding input creation operator given by $\hat{a}_i^\dagger$. The scattering matrix (see Eq. 3.7) allows us to express these input operators in terms of the output operators $\hat{b}_j^\dagger = \mathscr{S}_{i,j}\hat{a}_i^\dagger$. As suggested in

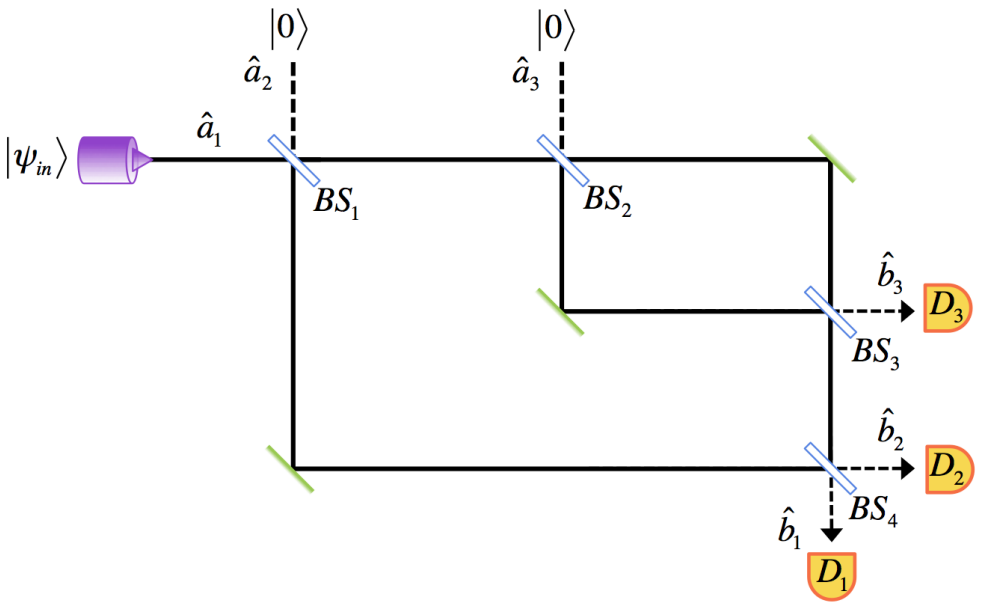


Figure 3.3 Optical diagram of an NMZI device. There are six possible detection outcomes (three spatial outcomes, with horizontal or vertical polarization).

the previous section, we shall define our polarization axis such that the input states always are of horizontal polarization. We also restrict our input states to initially occupy the first spatial input port, such that $|\psi_{in}\rangle = \hat{a}^{\dagger}_{1,H}|0\rangle$.

The two nested beam-splitters of an NMZI have their reflection and transmission coefficients set such that $r_2 = t_2 = r_3 = t_3 = \frac{1}{\sqrt{2}}$, and the scattering matrix (with $\hat{f}^{(k)}(\theta) = \hat{1}$ in Eq. 3.7) is given by

$$\mathscr{S}_{NMZI} = \begin{pmatrix} r_1 r_4 & t_1 r_4 & t_4 & 0 & 0 & 0 \\ -r_1 t_4 & -t_1 t_4 & r_4 & 0 & 0 & 0 \\ t_1 & -r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_1 r_4 & t_1 r_4 & t_4 \\ 0 & 0 & 0 & -r_1 t_4 & -t_1 t_4 & r_4 \\ 0 & 0 & 0 & t_1 & -r_1 & 0 \end{pmatrix}. \tag{3.30}$$

For our input state, $|\psi_{in}\rangle$, the output probabilities are given by

$$P(n_{1,H} = 1|\theta) = r_1^2 r_4^2, \tag{3.31}$$
$$P(n_{1,V} = 1|\theta) = 0, \tag{3.32}$$
$$P(n_{2,H} = 1|\theta) = r_1^2 t_4^2, \tag{3.33}$$
$$P(n_{2,V} = 1|\theta) = 0, \tag{3.34}$$

$$P(n_{3,H}=1|\theta)=t_1^2, \tag{3.35}$$
$$P(n_{3,V}=1|\theta)=0. \tag{3.36}$$

This device can be used for parameter estimation by placing the interaction of interest in one of five different locations of the NMZI. (See Fig. 3.4). Let us now proceed with an evaluation of these scenarios.

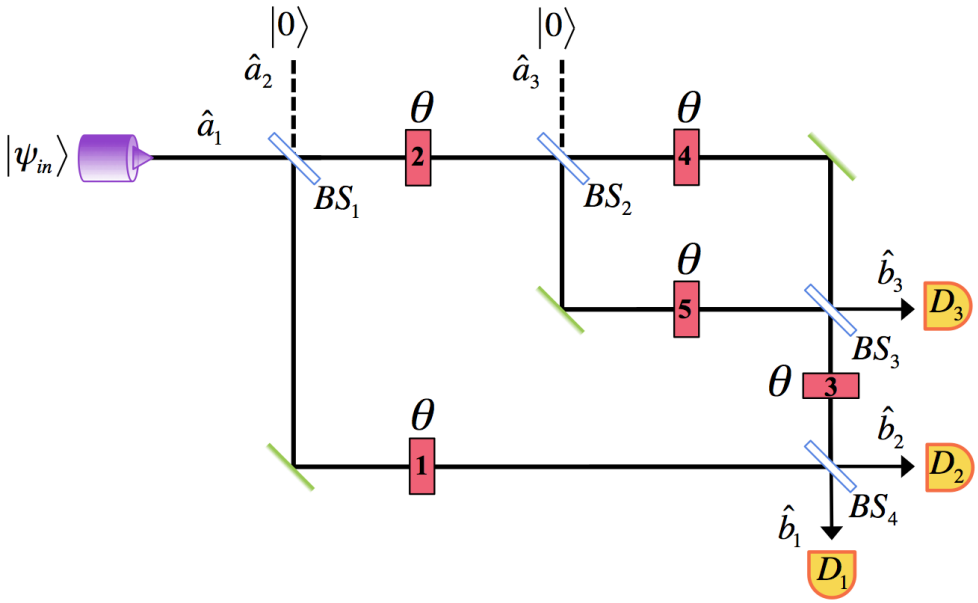


Figure 3.4 A polarization rotation is inserted in one, but only one, of the positions, 1-5, of the NMZI from Fig. 3.3.

One

In this first scenario we insert the polarizing rotator in the lower path between the first and fourth beam-splitter (see Fig. 3.5).

The scattering matrix of this scenario is given by

$$\mathscr{S}_1=\begin{pmatrix} r_1r_4 2\overline{\theta} & t_1r_4 2\overline{\theta} & t_4 & r_1r_4\theta & t_1r_4\theta & 0 \\ -r_1t_4 2\overline{\theta} & -t_1t_4 2\overline{\theta} & r_4 & -r_1t_4\theta & -t_1t_4\theta & 0 \\ t_1 & -r_1 & 0 & 0 & 0 & 0 \\ -r_1r_4\theta & -t_1r_4\theta & 0 & r_1r_4 2\overline{\theta} & t_1r_4 2\overline{\theta} & t_4 \\ r_1t_4\theta & t_1t_4\theta & 0 & -r_1t_4 2\overline{\theta} & -t_1t_4 2\overline{\theta} & r_4 \\ 0 & 0 & 0 & t_1 & -r_1 & 0 \end{pmatrix}, \tag{3.37}$$

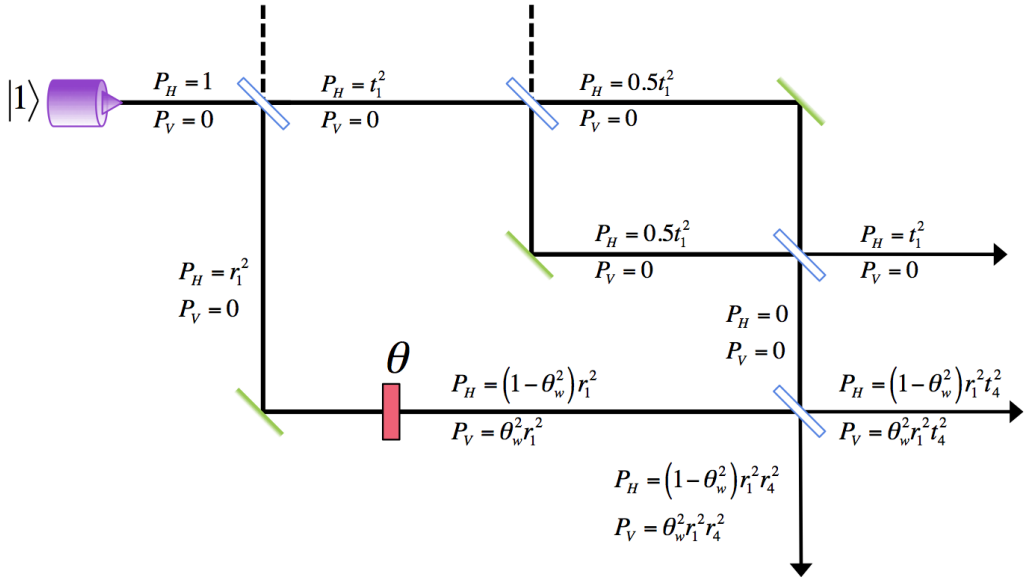


Figure 3.5 The probability density as the wavefunction propagates through the NMZI device with the polarization rotator placed in the lower path between BS_1 and BS_4.

where $\overline{\theta} \equiv \sqrt{1-\theta^2}/2$. Figure 3.5 shows the evolution of the probability density of the specific paths of the NMZI, assuming the input state, $|\psi_{in}\rangle$, from before.

We obtain the following conditional probabilities for the different output measurement events:

$$P(n_{1,H} = 1|\theta) = r_1^2 r_4^2 (1-\theta^2), \tag{3.38}$$
$$P(n_{1,V} = 1|\theta) = r_1^2 r_4^2 \theta^2, \tag{3.39}$$
$$P(n_{2,H} = 1|\theta) = r_1^2 t_4^2 (1-\theta^2), \tag{3.40}$$
$$P(n_{2,V} = 1|\theta) = r_1^2 t_4^2 \theta^2, \tag{3.41}$$
$$P(n_{3,H} = 1|\theta) = t_1^2, \tag{3.42}$$
$$P(n_{3,V} = 1|\theta) = 0. \tag{3.43}$$

The Shannon mutual information is given by Eq. 3.1:

$$H(\theta : \mathscr{M}) = \frac{\ln(108) - 4}{3\ln(2)} r_1^2. \tag{3.44}$$

Furthermore, the Fisher information of this device is calculated by either Eq. 3.2 or Eq. 3.21:

$$F = \frac{4}{1-\theta^2} r_1^2. \tag{3.45}$$

If we compare the information in this scenario with that of the free-space interaction considered above, we see that that both the Shannon mutual information and the classical Fisher information are scaled by the reflection coefficient, r_1, of the first beam-splitter (BS_1). The information is reduced exactly by the probability density that is scattered through parts of the NMZI that do not pass through the polarization rotator in the Schrödinger picture.

Two

We can now move to our next scenario, where we place the polarizing rotator in the upper path of the interferometer, between the first and second beam-splitter. In this scenario, the scattering matrix takes the form of

$$\mathscr{S}_2 = \begin{pmatrix} r_1 r_4 & t_1 r_4 & t_4 & 0 & 0 & 0 \\ -r_1 t_4 & -t_1 t_4 & r_4 & 0 & 0 & 0 \\ t_1 2\overline{\theta} & -r_1 2\overline{\theta} & 0 & t_1\theta & -r_1\theta & 0 \\ 0 & 0 & 0 & r_1 r_4 & t_1 r_4 & t_4 \\ 0 & 0 & 0 & -r_1 t_4 & -t_1 t_4 & r_4 \\ -t_1\theta & r_1\theta & 0 & t_1 2\overline{\theta} & -r_1 2\overline{\theta} & 0 \end{pmatrix}. \tag{3.46}$$

Figure 3.6 shows the evolution of the probability density as the wavepacket evolves through the NMZI.

The conditional probabilities for the different measurement outcomes are given by

$$P(n_1^H = 1|\theta) = r_1^2 r_4^2, \tag{3.47}$$
$$P(n_1^V = 1|\theta) = 0, \tag{3.48}$$
$$P(n_2^H = 1|\theta) = r_1^2 t_4^2, \tag{3.49}$$
$$P(n_2^V = 1|\theta) = 0, \tag{3.50}$$
$$P(n_3^H = 1|\theta) = t_1^2\left(1-\theta^2\right), \tag{3.51}$$
$$P(n_3^V = 1|\theta) = t_1^2\theta^2. \tag{3.52}$$

We can see many similarities between these probabilities to those in Scenario One. However, the θ-dependence has now been transferred from the first and second output paths to the third.

The Shannon mutual information is given by the expression

$$H(\theta : \mathscr{M}) = \frac{\ln(108) - 4}{3\ln(2)} t_1^2. \tag{3.53}$$

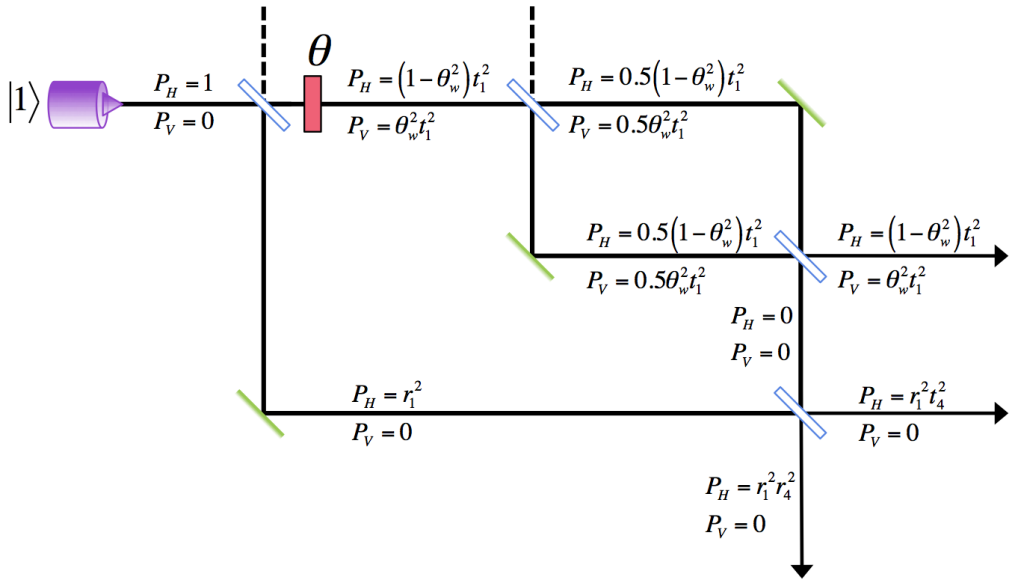


Figure 3.6 The probability density as the wavefunction propagates through the NMZI device with the polarization rotator placed in the upper path between BS_1 and BS_2.

Additionally, the Fisher information is scaled similarly and is given by

$$F = \frac{4}{1-\theta^2} t_1^2. \tag{3.54}$$

Comparing these results with Scenario One, we see how the previous r_1 dependency has been transformed to a t_1 dependency. Owing to the structure of the NMZI, the inner MZI has no effect on the information measures if the polarization rotator is placed just after the first beam-splitter.

Three

We continue our analysis of the NMZI by investigating how the aforementioned properties change if the polarizing rotator is placed between the third and the fourth beam-splitter. The scattering matrix is now given by

$$\mathscr{S}_3 = \begin{pmatrix} r_1 r_4 & t_1 r_4 & t_4 2\overline{\theta} & 0 & 0 & t_4\theta \\ -r_1 t_4 & -t_1 t_4 & r_4 2\overline{\theta} & 0 & 0 & r_4\theta \\ t_1 & -r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -t_4\theta & r_1 r_4 & t_1 r_4 & t_4 2\overline{\theta} \\ 0 & 0 & -r_4\theta & -r_1 t_4 & -t_1 t_4 & r_4 2\overline{\theta} \\ 0 & 0 & 0 & t_1 & -r_1 & 0 \end{pmatrix}. \tag{3.55}$$

In this third scenario, the probability densities are independent of θ, as can be seen in Fig. 3.7.

The conditional probabilities are given by

$$P(n_1^H = 1|\theta) = r_1^2 r_4^2, \tag{3.56}$$
$$P(n_1^V = 1|\theta) = 0, \tag{3.57}$$
$$P(n_2^H = 1|\theta) = r_1^2 t_4^2, \tag{3.58}$$
$$P(n_2^V = 1|\theta) = 0, \tag{3.59}$$
$$P(n_3^H = 1|\theta) = t_1^2, \tag{3.60}$$
$$P(n_3^V = 1|\theta) = 0. \tag{3.61}$$

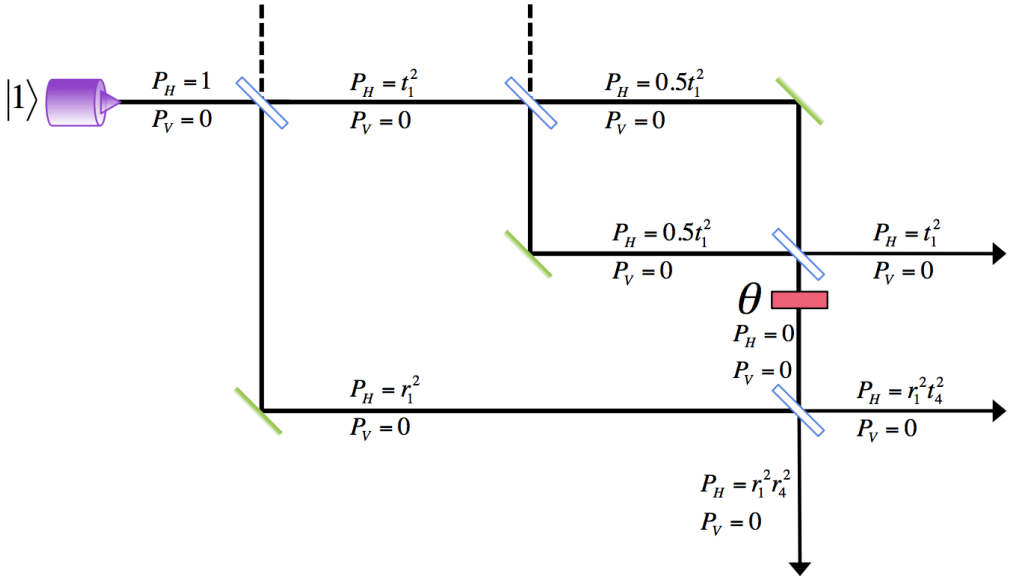


Figure 3.7 The probability density as the wavefunction propagates through the NMZI device with the polarization rotator placed between BS_3 and BS_4.

The fact that these probabilities are independent of θ is explained by the destructive interference of the inner MZI, which prohibits any part of the wavepacket from propagating through the spatial location of the rotator in this scenario.

As θ does not occur in any of the output probabilities, both the Shannon mutual information and the classical Fisher information will be zero:

$$H(\theta : \mathscr{M}) = 0, \tag{3.62}$$

$$F = 0. \tag{3.63}$$

Four and Five

To finalise the analysis of parameter estimation with NMZIs, we consider the placement of the polarization rotator within the inner MZI. Scenario Four and Scenario Five have the polarizing rotator placed in the upper and lower part of the inner MZI, respectively. Their corresponding scattering matrices are given by

$$\mathscr{S}_4 = \begin{pmatrix} r_1 r_4 - t_1 t_4 \vartheta^- & t_1 r_4 + r_1 t_4 \vartheta^- & t_4 \vartheta^+ & t_1 t_4 \overline{\vartheta} & -r_1 t_4 \overline{\vartheta} & t_4 \overline{\vartheta}^2 \\ -r_1 t_4 - t_1 r_4 \vartheta^- & -t_1 t_4 + r_1 r_4 \vartheta^- & r_4 \vartheta^+ & t_1 r_4 \overline{\vartheta} & -r_1 r_4 \overline{\vartheta} & r_4 \overline{\vartheta} \\ t_1 \vartheta^+ & -r_1 \vartheta^+ & -\vartheta^- & t_1 \overline{\vartheta} & -r_1 \overline{\vartheta} & \overline{\vartheta} \\ -t_1 t_4 \overline{\vartheta} & r_1 t_4 \overline{\vartheta} & -t_4 \overline{\vartheta} & r_1 r_4 - t_1 t_4 \vartheta^- & t_1 r_4 + r_1 t_4 \vartheta^- & t_4 \vartheta^+ \\ -t_1 r_4 \overline{\vartheta} & r_1 r_4 \overline{\vartheta} & -r_4 \overline{\vartheta} & -r_1 t_4 - t_1 r_4 \vartheta^- & -t_1 t_4 + r_1 r_4 \vartheta^- & r_4 \vartheta^+ \\ -t_1 \overline{\vartheta} & r_1 \overline{\vartheta} & -\overline{\vartheta} & t_1 \vartheta^+ & -r_1 \vartheta^+ & -\vartheta^- \end{pmatrix}, \tag{3.64}$$

$$\mathscr{S}_5 = \begin{pmatrix} r_1 r_4 - t_1 t_4 \vartheta^- & t_1 r_4 + r_1 t_4 \vartheta^- & -t_4 \vartheta^+ & -t_1 t_4 \overline{\vartheta} & r_1 t_4 \overline{\vartheta} & t_4 \overline{\vartheta} \\ -r_1 t_4 - t_1 r_4 \vartheta^- & -t_1 t_4 + r_1 r_4 \vartheta^- & -r_4 \vartheta^+ & -t_1 r_4 \overline{\vartheta} & r_1 r_4 \overline{\vartheta} & r_4 \overline{\vartheta} \\ t_1 \vartheta^+ & -r_1 \vartheta^+ & \vartheta^- & -t_1 \overline{\vartheta} & r_1 \overline{\vartheta} & \overline{\vartheta} \\ -t_1 t_4 \overline{\vartheta} & r_1 t_4 \overline{\vartheta} & t_4 \overline{\vartheta} & r_1 r_4 - t_1 t_4 \vartheta^+ & t_1 r_4 + r_1 t_4 \vartheta^+ & -t_4 \vartheta^- \\ -t_1 r_4 \overline{\vartheta} & r_1 r_4 \overline{\vartheta} & r_4 \overline{\vartheta} & -r_1 t_4 - t_1 r_4 \vartheta^+ & -t_1 t_4 + r_1 r_4 \vartheta^+ & -r_4 \vartheta^- \\ -t_1 \overline{\vartheta} & r_1 \overline{\vartheta} & \overline{\vartheta} & t_1 \vartheta^- & -r_1 \vartheta^- & \vartheta^+ \end{pmatrix}, \tag{3.65}$$

where $\vartheta^\pm \equiv (1 \pm \vartheta)/2$ and where a temporary and superficial change of variables has been made, such that $\vartheta \equiv \sqrt{1-\theta^2}$. Figures 3.8 and 3.9 show the corresponding evolutions of the probability density of Scenario Four and Scenario Five, respectively.

Compared to scenarios One, Two and Three, the conditional probabilities of the measurement outcomes now take more complicated forms. For Scenario Four, they are given by

$$P(n_1^H = 1|\vartheta_w) = \frac{1}{4}[2r_1 r_4 - t_1 t_4(1-\vartheta_w)]^2, \tag{3.66}$$

$$P(n_1^V = 1|\vartheta_w) = \frac{1}{4} t_1^2 t_4^2 (1-\vartheta_w^2), \tag{3.67}$$

$$P(n_2^H = 1|\vartheta_w) = \frac{1}{4}[2r_1 t_4 + t_1 r_4(1-\vartheta_w)]^2, \tag{3.68}$$

$$P(n_2^V = 1|\vartheta_w) = \frac{1}{4} t_1^2 r_4^2 (1-\vartheta_w^2), \tag{3.69}$$

$$P(n_3^H = 1|\vartheta_w) = \frac{1}{4} t_1^2 (1+\vartheta_w)^2, \tag{3.70}$$

$$P(n_3^V = 1|\vartheta_w) = \frac{1}{4} t_1^2 (1-\vartheta_w^2). \tag{3.71}$$

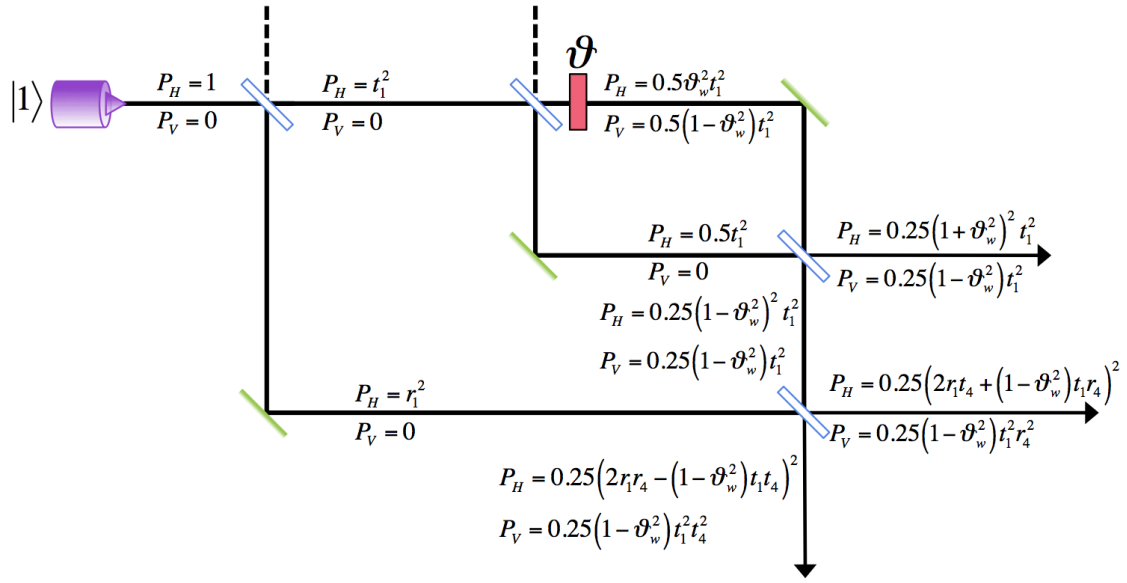


Figure 3.8 The probability density as the wavefunction propagates through the NMZI device with the polarization rotator placed in the upper path of the inner MZI, between BS_2 and BS_3.

These probabilities take the same forms in Scenario Five, with the exception of two sign changes such that $P(n_1^H = 1|\vartheta_w) = \frac{1}{4}[2r_1r_4 + t_1t_4(1-\vartheta_w)]^2$ and $P(n_2^H = 1|\vartheta_w) = \frac{1}{4}[2r_1t_4 - t_1r_4(1-\vartheta_w)]^2$.

The corresponding classical Fisher information for the devices shown in Figs. 3.8 and 3.9 is given by

$$F(\theta) = \frac{2}{1-\theta^2}t_1^2. \tag{3.72}$$

In accordance with the discussion in Section 4 and Eq. 3.21, the Fisher information is directly proportional to the integrated probability density that passes through the polarization rotator in the Schrödinger picture. If we compare the scenarios in this section with Scenario Two, the probability density at the rotator has now been halved by the second beam-splitter. Hence, the Fisher information in Eq. 3.72 is halved compared to its value in Scenario Two (Eq. 3.54).

Throughout the five scenarios of this section, we have studied the simplicity of the Fisher information—confirming Eq. 3.21. However, in scenarios One, Two and Three, the Shannon mutual information has been directly proportional to the Fisher information. This rather beautiful relationship is broken when studying the Shannon mutual information of scenarios Four and Five. The corresponding Shannon mutual information for these scenarios takes a complicated form. However, by making the assumption that $t_4 = r_1$ and $r_4 = t_1$, it is possible

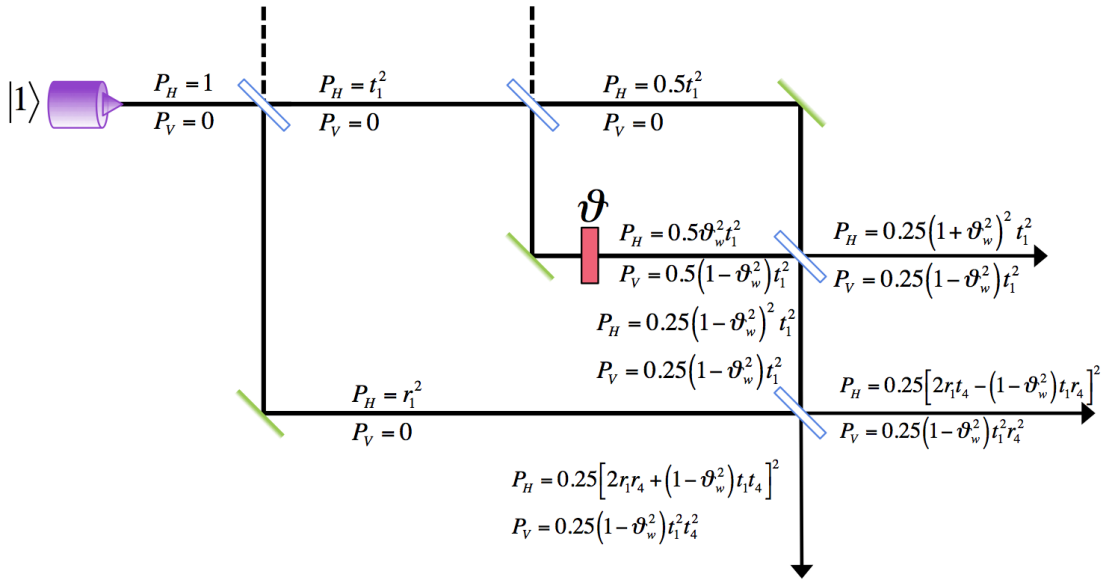


Figure 3.9 The probability density as the wavefunction propagates through the NMZI device with the polarization rotator placed in the lower path of the inner MZI, between BS_2 and BS_4.

to somewhat simplify the expression such that

$$H(\vartheta:\mathscr{M}) = \frac{1}{3\ln(2)t_1^2}\Big[-2r_1^3\ln(r_1^2)+t_1^2\{3\ln(3)+t_1^2[\ln(2)-1] \\ -2-(3r_1^2+t_1^4)\ln(3r_1^2+t_1^4)\}\Big]. \tag{3.73}$$

We can simplify the Shannon mutual information of Eq. 3.73 further by employing approximations. For example, if $t_1 \approx 0$, the Shannon mutual information is well modelled by a second-order term:

$$H(\vartheta:\mathscr{M}) \approx \frac{-3+\ln(2)+3\ln(3)}{3\ln(2)}t_1^2. \tag{3.74}$$

Table 3.1 Values of the numerical constants in Eq. 3.75, the [6/4] Padé approximation of Eq. 3.73.

i	a_i	b_i
2	$\frac{-3+\ln(2)+3\ln(3)}{3\ln(2)}$	-1
4	$\frac{25-6\ln(2)+25\ln(3)}{18\ln(2)}$	$\frac{-1}{10[-7+3\ln(3)]}$
6	$\frac{254-3\ln(2)-429\ln(3)+180\ln(3)^2}{90[-7+3\ln(3)]\ln(3)}$	

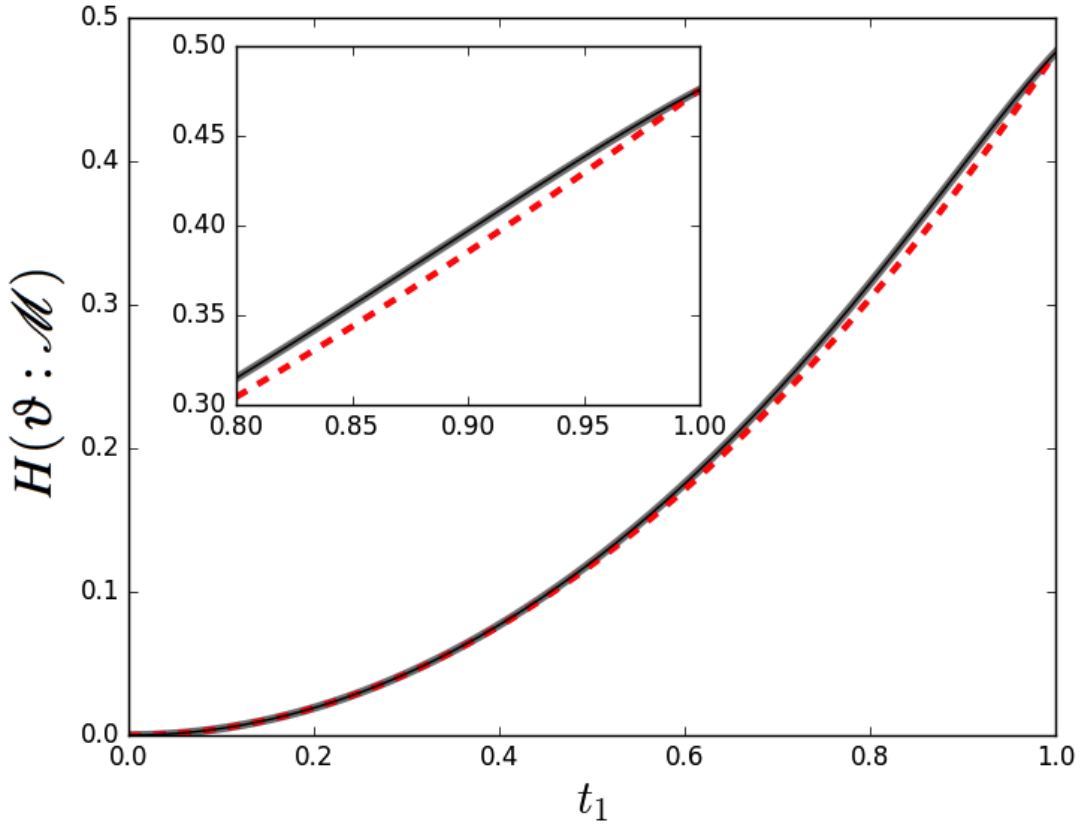


Figure 3.10 The Shannon mutual information between the polarization rotation parameter, ϑ, and the measurement outcomes, $\{\mathscr{M}_i\}$, as a function of beam-splitter transmission coefficient, t_1, as described in the main text. The solid black line shows the true curve of Eq. 3.73, the thick grey line shows the Padé approximation (virtually indistinguishable from the true curve), and the red dashed line shows the second order Taylor expansion. This figure originally appeared in Ref. [47].

If we want a simplified model that is accurate for all values of t_1 ($0 \leq t_1 \leq 1$), we can make use of a Padé approximant (see Ref. [109]) of order $[6/4]$:

$$H(\vartheta:\mathscr{M}) \approx \frac{a_2 t_1^2 + a_4 t_1^4 + a_6 t_1^6}{1 + b_2 t_1^2 + b_4 t_1^4}, \tag{3.75}$$

where the constants a_i and b_j for $i \in \{2,4,6\}$ and $j \in \{2,4\}$ are presented in Table 3.1.

Figure 3.10 shows the Shannon mutual information from Eq. 3.73 and the two approximate models as functions of t_1. The second order Taylor expansion (for $t_1 \leq 0.4$) and the full Padé approximation model the true curve within mean squared errors of 3.1×10^{-8} and 2.8×10^{-9}, respectively.

After having worked our way through this technical study of parameter estimation with NMZI structures, we proceed by investigating how the Fisher information measure can be used to scrutinise type I counterfactuality of protocols, which are based on these structures.

6 Measures of Counterfactual Violations

Counterfactual interaction-free phenomena are processes with outcomes that classically only happen if particles take certain paths, but that—owing to quantum mechanics—can happen anyway, without these particle paths: i.e. *counterfactually*. However, as mentioned before, there are conceptual problems with assigning past paths to quantum particles. What might seem a plausible methodology for one scenario is implausible for another. In this section we shall see how concepts of quantum probabilities and the classical Fisher information can be used to design two operational and interpretation-independent measures for counterfactual violations in the type I and type II protocols. We shall then use these measures to scrutinise the protocols.

6.1 Differences between type I and type II counterfactuality

In Chapter 2, we studied the type I and type II CFC protocols. These protocols differ in both their operational setup and their definition of counterfactuality. Hence, they ought to be evaluated using different measures of counterfactuality. Before we develop these measures, let me present, once again, the respective interferometers of the type I (Fig. 3.11) and the type II (Fig. 3.12) protocols. Table 3.2 gives a brief summary of the spatial evolution of the particles in these two types of CFC.

Table 3.2 The differences in the particle propagation between type I and type II CFC as given in Ref. [47]. In both protocols Alice produces the particles, and Bob transmits a message.

	Logical 0	Logical 1
Type I:	No particles cross the transmission line between Alice and Bob.	
Type II:	Particles propagate from Alice to Bob via the transmission line.	Particles propagate from Alice to the transmission line and back again.

In both schemes, the processes with inserted objects in Bob's laboratory (logical 1) satisfies their respective definitions. This is simply because if the quantum Zeno effect is induced by "blocking" Bob's laboratory, it is impossible for the "surviving" wavefunction to

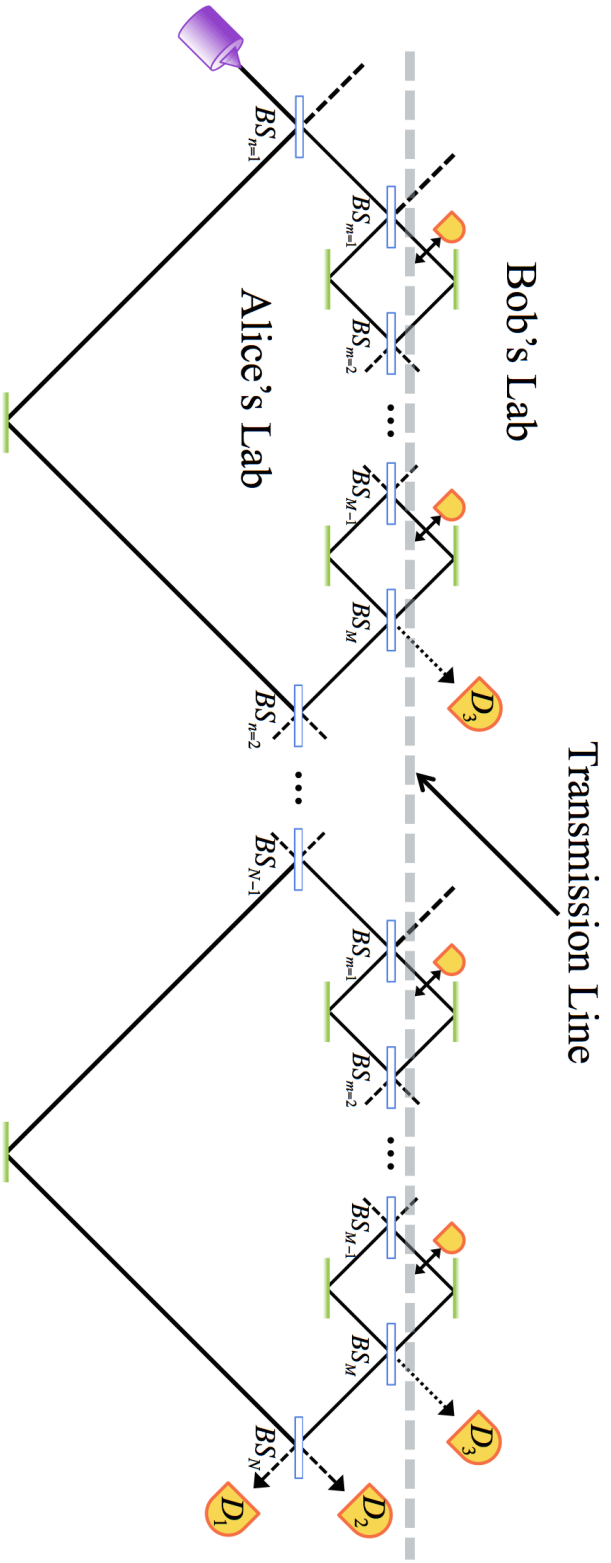


Figure 3.11 The CNMZI device suggested for type I CFC by Salih *et al*. in Ref. [39]. Alice sends particles to the input path. Bob can control the output statistics at D_1 and D_2 by inserting, or not inserting, detectors in his laboratory.

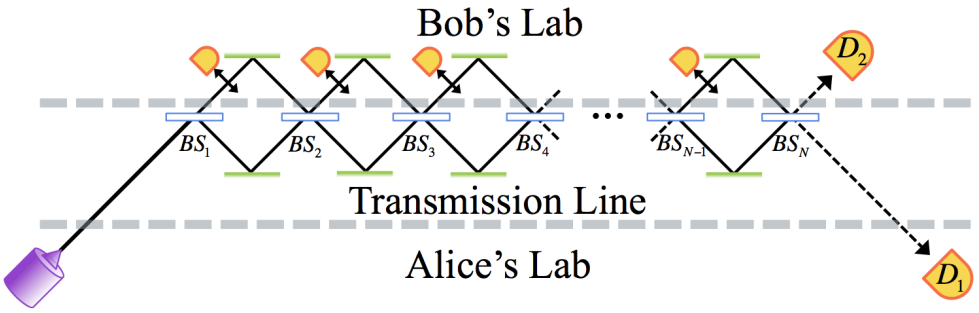


Figure 3.12 The CMZI device suggested for type II CFC by Crispin Barnes and myself in Ref. [46] (see also Fig. 2.8). Alice sends particles to the input path, and Bob decides the output statistics at D_1 and D_2 by inserting, or not inserting, detectors in his laboratory.

have interacted with Bob. Consequently, there is no controversy regarding the counterfactual nature of these processes. This leaves us with the task of developing a methodology for the evaluation of the level of counterfactuality of the logical 0 processes in the respective types of CFC.

6.2 Measure of type I counterfactuality

The type I definition of counterfactuality aims to establish communication without any particles travelling from Bob to Alice, and vice versa. Moreover, all particles in both types of CFC originate from Alice's laboratory. Hence, a violation of type I counterfactuality would occur if a quantum particle were to travel from Alice to Bob. To complicate the investigation further, in the type I protocol all particles also finish in Alice's laboratory. Hence, to establish if these particles were ever present in Bob's laboratory one needs a framework to determine inter-measurement presence. This is where the analysis of Section 4 comes in handy. Given the discussion regarding the classical Fisher information in optical circuits, it is evident that a good measure of a counterfactual violation can be based on the Fisher information encoded in Bob's laboratory, in a particle that originates from Alice's. I shall thus investigate counterfactual violations of a type I logical 0 process by considering the tagging of wavepackets within Bob's laboratory.

However, in order to establish what the "strength" of a counterfactual violation is, we need a Fisher information benchmark. It seems suitable to set this benchmark such that it corresponds to the classical Fisher information obtained from a particle that originates from Bob, where it directly interacts with a tagging medium and propagates, in free space, to Alice. Such a scenario would clearly violate counterfactuality. We already calculated this quantity, F_{free}, in Section 5.1 (see Eq. 3.27). This allows us to define a measure for type I counterfactual violations. I call this measure the type I counterfactual violation strength:

$$\mathscr{D} := \frac{F}{F_{free}}. \tag{3.76}$$

The value of $\mathscr{D}$ can effectively be thought of as the extent that a particle, originating from Alice, has an inter-measurement presence in Bob's laboratory, owing to some weak disturbance or interaction there, in units of the Fisher information of a free-space interaction. Hence, a perfectly counterfactual type I process would generate a value of $\mathscr{D} = 0$, confirming that no wavefunction has leaked into Bob's laboratory. A process in which $\mathscr{D} \geq 1$, however, will have violated counterfactuality to a degree stronger than or equal to a free-space transmission of a particle between Bob and Alice. Consequently, protocols in which $\mathscr{D}$ is greater than unity are convicted of fully violating type I counterfactuality.

6.3 Measure of type II counterfactuality

The evaluation of counterfactual violations of type II processes requires a different measure to Eq. 3.76. The analysis is simpler, at least from a conceptual point of view. This is because type II protocols allow particles that originate from Alice's laboratory to be encoded in Bob's laboratory, as long as they do not return to her. The probability of detecting a particle in Alice's laboratory in a logical 0 process is null for vanishingly weak tagging interactions ($\theta = 0$). Hence, we can define a measure of type II counterfactual violations with respect to the total probability of observing that a particle returns to Alice as a consequence of a weak, non-collapsing, disturbance ($\theta \neq 0$) in Bob's laboratory. In order to adequately define this measure I need to introduce some notation. Let me denote the measurement outcomes triggered by a particle detection in the Hilbert space of states within the spatial extent of $\boldsymbol{A}$ (Alice) by: $\mathscr{M}_{j\in J_{\boldsymbol{A}}}$. Additionally, let me denote the negative measurement that indicate all outcome states outside the spatial extent of $\boldsymbol{A}$ by: $\mathscr{M}_{j'\notin J_{\boldsymbol{A}}}$. The measure of type II counterfactual violation strengths can then be expressed as

$$P_{\boldsymbol{A}} := \sum_{j\in J_{\boldsymbol{A}}} P(\mathscr{M}_j|\theta,\psi_{in}) = 1 - P(M_{j'\notin J_{\boldsymbol{A}}}|\theta,\psi_{in}). \tag{3.77}$$

This measure is such that a maximum violation of counterfactuality in a type II logical 0 process would generate a value of $P_{\boldsymbol{A}} = 1$. A perfect process, on the other hand, would mean that $P_{\boldsymbol{A}} = 0$.

Consider a process where the probability of the particle being detected in Alice's laboratory is very small, such that $P_{\boldsymbol{A}} \approx 0$. The nature of the classical Fisher information, Eq. 3.2, is such that even small probabilities can generate large Fisher information components. Consequently, even in scenarios where the probability of Alice detecting particles is very low,

she can still gain a relatively large amount of information of Bob's laboratory. The amount of Fisher information acquired by an observer, $\boldsymbol{A}$, in a given experiment can thus be defined as a spatially conditioned version of Eq. 3.2:

$$F_{\boldsymbol{A}}(\theta) := \sum_{j \in J_{\boldsymbol{A}}} \frac{1}{P(\mathscr{M}_j|\theta,\psi_{in})} \left[\frac{\partial}{\partial\theta} P(\mathscr{M}_j|\theta,\psi_{in})\right]^2 + \frac{1}{1-P_{\boldsymbol{A}}} \left[\frac{\partial}{\partial\theta} P_{\boldsymbol{A}}\right]^2. \tag{3.78}$$

The first term of the formula gives the Fisher information contribution from the particle detections made by observer $\boldsymbol{A}$. The second term gives the contribution resulting from the negative measurements of $\boldsymbol{A}$, i.e. in scenarios where no particle is detected by $\boldsymbol{A}$. Interestingly, in a Type II scheme, θ and $P_{\boldsymbol{A}}$ can be small, such that counterfactuality only is weakly violated, at the same time as Alice obtains a large Fisher information of the tagging interaction in Bob's laboratory.

7 Evaluation of Type I Counterfactuality

The type I CFC scheme of Salih *et al.* [39], has been outlined in Chapter 2 and discussed in Section 6.1 of this chapter. The protocol is based on a concatenated nested MZI (CNMZI) device. The bit transmission efficiencies of the two type I processes are shown in Fig. 3.13. The success probabilities of detections at D_1 (0-bit) or D_2 (1-bit) are calculated with the assumption that the quantum channels of the device and its unitary operations are perfect. Even in this unrealistic scenario, Fig. 3.13 (previously presented as Fig. 2.5 in Chapter 2.3) shows that in order to reduce bit errors it is required that $M \gg N$. If we wish to keep the bit errors of the individual processes below 95 %, Fig. 3.13(a) shows that we need $N \approx 50$ (for the logical 0). But from Fig. 3.13(b) we know that such a value of N then requires that $M \approx 1,200$ (for the logical 1). As a result, in order to limit the number of bit errors of the individual processes to less than 5 %, it is necessary to use a CNMZI device with at least $60,000$ perfect beam-splitters...

As a consequence of this, the type I CFC protocol naturally involves a large number of unitary transformations. This complicates the investigation of the protocol with respect to the tagging methodology.

7.1 A single NMZI

Given the large number of unitary operations involved in the full type I device, an analytical evaluation is tedious. In Refs. [40, 41, 79] it has been suggested that by considering pre-

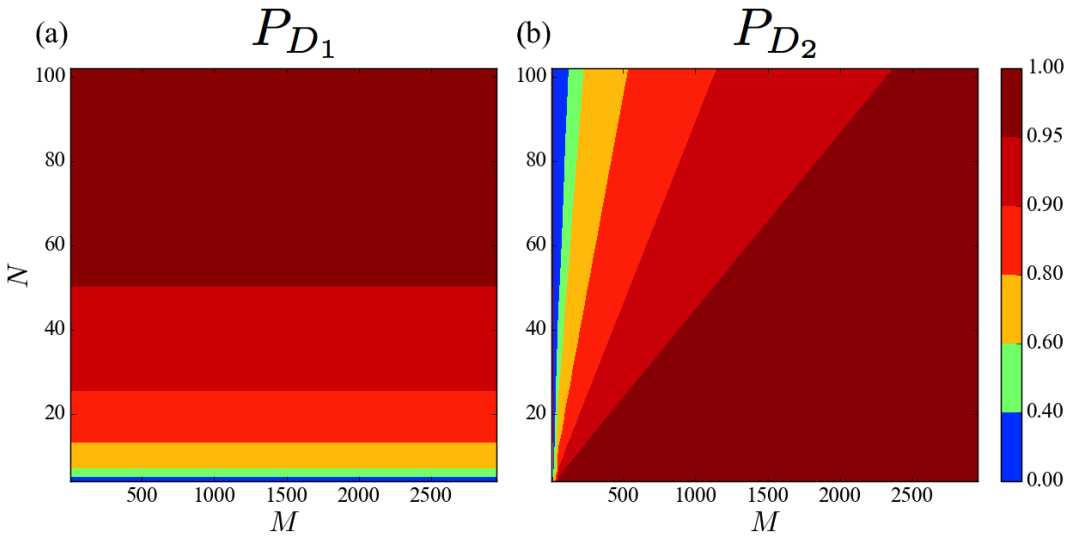


Figure 3.13 The success probability of the (a) logical 0 and (b) logical 1 process of the type I CFC scheme. This figure originally appeared in Ref. [47] and is in agreement with Ref. [39].

and post-selected events, the conceptual problem of the concatenated NMZI used in the type I protocol can be reduced to a study of a single NMZI device (see Chapter 2.3.2). These references use this reduction to present analytical arguments regarding the counterfactuality of the type I protocol. Whilst the suggestion of a reduction is completely plausible, these previous works fail to provide rigorous arguments as to why their specific reduction is a logical one. Crucially, the reduction that they consider is that of a single evaluation of an NMZI device, which does not allow for the transmission of logical bits. Below I shall present an alternative approach.

Theoretical analysis

Instead of theoretically analysing the full protocol of Salih *et al.*, I shall design an alternative post-selected type I protocol, which actually allows for communication. We can then evaluate the level to which this protocol satisfies type I counterfactuality.

Consider Fig. 3.14, which shows the spatial domains of the transmission line as well as Alice's and Bob's laboratories, distributed over a single NMZI. Let us require a pre-selection of our input states such that they take the form of $|\psi_{in}\rangle = a^{\dagger}_{1,H}|0\rangle$ (used previously in Section 5). Alice inputs these states with a pre-determined frequency. As in the previously described schemes, Bob then has the possibility of introducing an absorbing object (detector D_B) in his laboratory, or of keeping it open. However, now we also introduce a weak polarization rotator in Bob's laboratory, which mimics disturbances within his part of the device. Moreover, let

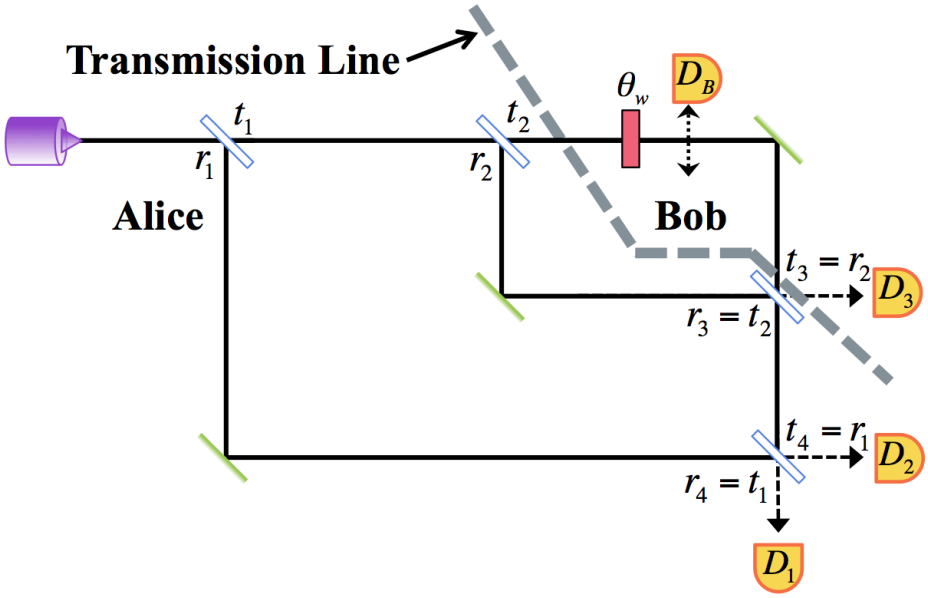


Figure 3.14 The NMZI as used in the communication scheme described in the main text.

us post-select our output states such that we do not consider any events where the photon is output to D_B or D_3 (such events would clearly violate type I counterfactuality).

Let me highlight a difference of this model compared with the post-selected reduction previously outlined in Chapter 2.3.2. In the previous weak trace analysis the post-selection was done with respect to only one output path [40, 41, 79]. However, there are two output paths within Alice's laboratory (see Fig. 3.14). Only in the ideal limit of an infinite number of beam-splitters does the 0-bit process generate a definite outcome. Hence, the reduced scheme can result in the photon being output to Alice laboratory from any of the two output paths. It is thus difficult to motivate the use of post-selection to ignore one of these output paths when discussing the past of a particle propagating through the device to Alice. That is why I consider both output paths.

If Bob wishes to transmit a logical 1 to Alice, he introduces his absorbing detector, D_B, into the particle path of his laboratory. After the output probabilities have been renormalised with respect to the post-selection criteria, they are given by

$$P_1(n_1^H = 1 | n_{3,B}^{H,V} = 0) = \mathscr{N}_1 \left(r_1 r_4 - r_2 t_1 t_2 t_4 \right)^2, \tag{3.79}$$

$$P_1(n_1^V = 1 | n_{3,B}^{H,V} = 0) = 0, \tag{3.80}$$

$$P_1(n_2^H = 1 | n_{3,B}^{H,V} = 0) = \mathscr{N}_1 \left(r_1 t_4 + r_2 r_4 t_1 t_2 \right)^2, \tag{3.81}$$

$$P_1(n_2^V = 1 | n_{3,B}^{H,V} = 0) = 0, \tag{3.82}$$

with $\mathscr{N}_1 \equiv \left[1 - t_1^2\left(r_2^4 + t_2^2\right)\right]^{-1}$.

We proceed by setting the reflection and transmission coefficients in accordance with Fig. 3.14 ($r_2 = t_2 = 1/\sqrt{2}$), and the probabilities simplify considerably:

$$P_1(n_1^H = 1 | n_{3,B}^{H,V} = 0) = \frac{r_1^2 t_1^2}{3r_1^2 + 1}, \tag{3.83}$$

$$P_1(n_1^V = 1 | n_{3,B}^{H,V} = 0) = 0, \tag{3.84}$$

$$P_1(n_2^H = 1 | n_{3,B}^{H,V} = 0) = \frac{\left(r_1^2 + 1\right)^2}{3r_1^2 + 1}, \tag{3.85}$$

$$P_1(n_2^V = 1 | n_{3,B}^{H,V} = 0) = 0. \tag{3.86}$$

As the output probabilities of the 1-bit process do not depend on θ_w, the corresponding Fisher information is naturally given by

$$F^1 = 0. \tag{3.87}$$

This is a direct consequence of the fact that any part of the wavefunction that interacts with the tagging mechanism in Bob's laboratory is absorbed by detector D_B and cannot return to Alice's laboratory.

Let us now study the logical 0 process instead, in which Bob keeps the path through his laboratory free from detectors. In this scenario the renormalised output probabilities are given by

$$P_0(n_1^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0\left(r_1 r_4 + r_2 t_1 t_2 t_4\left[2\overline{\theta}_w - 1\right]\right)^2, \tag{3.88}$$

$$P_0(n_1^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0\left(r_2^2 t_1^2 t_2^2 t_4^2 \theta_w^2\right), \tag{3.89}$$

$$P_0(n_2^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0\left(r_1 t_4 - r_2 r_4 t_1 t_2\left[2\overline{\theta}_w - 1\right]\right)^2, \tag{3.90}$$

$$P_0(n_2^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0\left(r_2^2 r_4^2 t_1^2 t_2^2 \theta_w^2\right), \tag{3.91}$$

with $\mathscr{N}_0 \equiv \left[1 - t_1^2\left(r_2^4 + t_2^4 + 4r_2^2 t_2^2 \overline{\theta}_w\right)\right]^{-1}$. We again simplify the probabilities by setting the reflection and transmission coefficients in accordance with Fig. 3.14 ($r_2 = t_2 = 1/\sqrt{2}$):

$$P_0(n_1^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0' r_1^2 t_1^2\left(1 + 2\overline{\theta}_w\right)^2, \tag{3.92}$$

$$P_0(n_1^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0' r_1^2 t_1^2 \theta_w^2, \tag{3.93}$$

$$P_0(n_2^H = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0'\left(1 + r_1^2 - 2t_1^2 \overline{\theta}_w\right)^2, \tag{3.94}$$

$$P_0(n_2^V = 1 | \theta_w, n_{3,B}^{H,V} = 0) = \mathscr{N}_0' t_1^4 \theta_w^2, \tag{3.95}$$

where $\mathscr{N}_0' \equiv \left[4 - 2t_1^2\left(1 + 2\overline{\theta}_w\right)\right]^{-1}$.

Now, using the probabilities for the 0-bit process we can calculate the corresponding Fisher information to be

$$F^0 = \frac{t_1^2}{1 - \theta_w^2}. \tag{3.96}$$

Note that whilst $F^1 = 0$ in the 1-bit process, such that no information of the parameter θ_w flows from Bob to Alice, in the 0-bit process the Fisher information takes a finite value. Also, note that even a vanishingly weak tagging polarization rotation would result in the transmission of information from Bob to Alice, as $\lim_{\theta_w \to 0} F^0 = t_1^2$.

As Bob has access to parts of the inner MZI, he is in control of the output probability distributions in Alice's laboratory in a similar manner to the Salih *et al.* scheme. However, in order for our setup to be representative of Salih's setup, which is based on a CNMZI, we require that $\theta_w \ll t_1 \ll r_1$. As a result of these restrictions, the polarization rotation of the wavefunction passing through the upper path of the inner MZI will have a minute impact on the output probability distributions. Moreover, as $t_1 \ll r_1$ the output statistics of the 0-bit and 1-bit process are very similar. In both bit processes Alice will measure the output particle in state $|n_2^H = 1\rangle$ with probabilities close to unity. So, how can Bob use the device from Fig. 3.14 and the two, above described, processes in order to communicate with Alice? Essentially, in order for Bob and Alice to communicate, they have to assign a large number of processes for each successful bit-transmission, such that Alice can decode each logical bit from the statistics of a longer bit string.

The communication scheme is as follows: Alice inputs a number, n_γ, of single particles (post-selected such that those collapsing on D_B or D_3 are excluded), one after the other, into the NMZI device of Fig. 3.14. If Bob wishes to transmit a logical 0 he keeps his laboratory open, and if he wishes to transmit a logical 1 he inserts detector D_B. Alice subsequently detects the n_γ particles and records the statistics of the various possible output states: $|n_1^H = 1\rangle$, $|n_2^H = 1\rangle$, $|n_1^V = 1\rangle$ or $|n_2^V = 1\rangle$. In the scenario where Alice measures a single or more events as $|n_1^V = 1\rangle$ or $|n_2^V = 1\rangle$, she will know with certainty that a logical 0 was sent. However, owing to the fact that θ_w is very small compared to t_1, the accumulative probability of these events happening is negligible. Consequently, Alice will—with high probability—have to make her prediction of the logical bit-value with regard to the statistics of the $|n_1^H = 1\rangle$ and $|n_2^H = 1\rangle$ events. This allows her to infer whether Bob has sent a logical 0 or a logical 1 from the number of particles, q, that she measures in the $|n_1^H = 1\rangle$ state.

As the total Fisher information increases linearly with the number of evaluations of an experiment, the question we are interested in answering is: What number of single-photon

evaluations, n_γ, of the NMZI device allows for an effective communication scheme with a limited number of bit errors?

Let us redefine $P_1 \equiv P_1(n_1^H = 1|n_{3,B}^{H,V} = 0)$ and $P_0 \equiv P_0(n_1^H = 1|\theta_w, n_{3,B}^{H,V} = 0)$. We note that $P_1 < P_0$ for $\theta_w \ll t_1 \ll 1$. Hence, Alice will record a 1-bit every time $q < q'$. In the limit of long message strings, we can assume that Bob will produce logical 0s and 1s at the same rates, and the exact value of q' is given by

$$q' = \left\lfloor \frac{n_\gamma \ln\left(\frac{1-P_1}{1-P_0}\right)}{\ln\left(\frac{P_0}{P_1}\right) - \ln\left(\frac{1-P_0}{1-P_1}\right)} \right\rfloor. \tag{3.97}$$

This formula can be used to evaluate the probability for a non-faulty logical bit-transmission:

$$P_{succ.} = \frac{1}{2}\sum_{q=0}^{q'} \frac{n_\gamma! P_1^q (1-P_1)^{n_\gamma - q}}{q!(n_\gamma - q)!} + \frac{1}{2}\sum_{q=q'+1}^{n_\gamma} \frac{n_\gamma! P_0^q (1-P_0)^{n_\gamma - q}}{q!(n_\gamma - q)!}. \tag{3.98}$$

Given the parameters of the setup, it is possible to find an acceptable value of n_γ by numerically solving Eq. 3.98. However, as already discussed, we require that the transmission coefficient is small [$t_1 = \sin(\pi/2N) \ll 1$], in order to make the investigation representative of the setup discussed in Ref. [36, 38–41, 43]. This allows us to simplify the calculation of n_γ. For $t_1 \ll 1$ we can—by the central limit theorem—assume that n_γ has to be large such that the two bit-processes will generate normally distributed events. These two processes will each have a mean situated at P_1 and P_0, respectively. Their respective standard deviations will be given by

$$\sigma_i = \sqrt{\frac{P_i(1-P_i)}{n_\gamma}}, \tag{3.99}$$

where $i = 0, 1$. We note that the standard deviations decrease reciprocally with the square-root of n_γ.

In order for Alice to be able to determine the value of the logical bit correctly with probability $1 - \varepsilon$ we require that

$$n_\gamma \geq \left(\Phi^{-1}(\varepsilon) \frac{\sqrt{P_0(1-P_0)} + \sqrt{P_1(1-P_1)}}{P_0 - P_1} \right)^2, \tag{3.100}$$

where $\Phi^{-1}(\varepsilon)$ is the inverse of the standard normal cumulative distribution function.

We continue our analysis by substituting the expressions for P_0 and P_1, and then Taylor expanding n_γ with respect to small values of t_1:

$$n_\gamma \geq \left(\Phi^{-1}(\varepsilon)\right)^2 \frac{4}{t_1^2} + \mathscr{O}(t_1^{-1}). \tag{3.101}$$

Notably, the number of evaluations needed for Alice to distinguish between the bit-processes sufficiently well is inversely proportional to the square of transmission coefficient t_1. Moreover, as the total Fisher information is proportional to n_γ (the number of processes per logical bit), the counterfactual violation strength of the type I logical 0 is given by

$$\mathscr{D} = n_\gamma \frac{F^0}{F_{free}} \gtrsim \left(\Phi^{-1}(\varepsilon)\right)^2. \tag{3.102}$$

If we require a success rate of roughly 95% ($\varepsilon = 0.05$), we obtain a value of $\mathscr{D} \approx 2.7$. Hence, we can conclude that even with post-selection of favourable measurement outcomes, the type I CFC scheme described in this section is less counterfactual than direct communication between Alice and Bob.

Numerical simulation of the particle propagation

The discussion of type I counterfactual protocols in this thesis has so far been based on the matrix representation of unitary operations. This is a great way of demonstrating theoretical arguments. However, any realisation of a quantum protocol will have to deal with wavepackets of finite (or—if you like—infinite) size, propagating through space. Remember that in Chapter 2.4.4, we used a toy model of a Gaussian particle in order to simulate what a more realistic evolution of a quantum particle looks like. Here, I shall use the same toolkit (outlined in Appendix A) to illustrate the origin of the counterfactual violations inside the NMZI device.

We simulate a massive Gaussian spin-$\frac{1}{2}$ particle that propagates through a device owing to the diabatic shifts of the corresponding minima of the external device potentials. For simplicity the NMZI has been mapped to a 1D structure. Furthermore, it is a formidable task to simulate NMZIs with $t_1 \approx 0$. Hence, we settle for a device with beam-splitter parameters set to $t_2 = r_2 = t_3 = r_3 = \frac{1}{\sqrt{2}}$, $t_1 = t_4 = \frac{1}{2}$ and $r_1 = r_4 = \frac{\sqrt{3}}{2}$. This simulation is presented in Fig. 3.15.

In Fig. 3.15(a) Alice inputs a single particle onto BS_1, contained within her laboratory. The external device potentials (green dashed lines) are then shifted such that the part of the wavefunction that was reflected is held at $x = 0.5$, whilst the transmitted part continues to propagate towards BS_2 [see Fig. 3.15(b)]. In Fig. 3.15(c) the propagating part of the

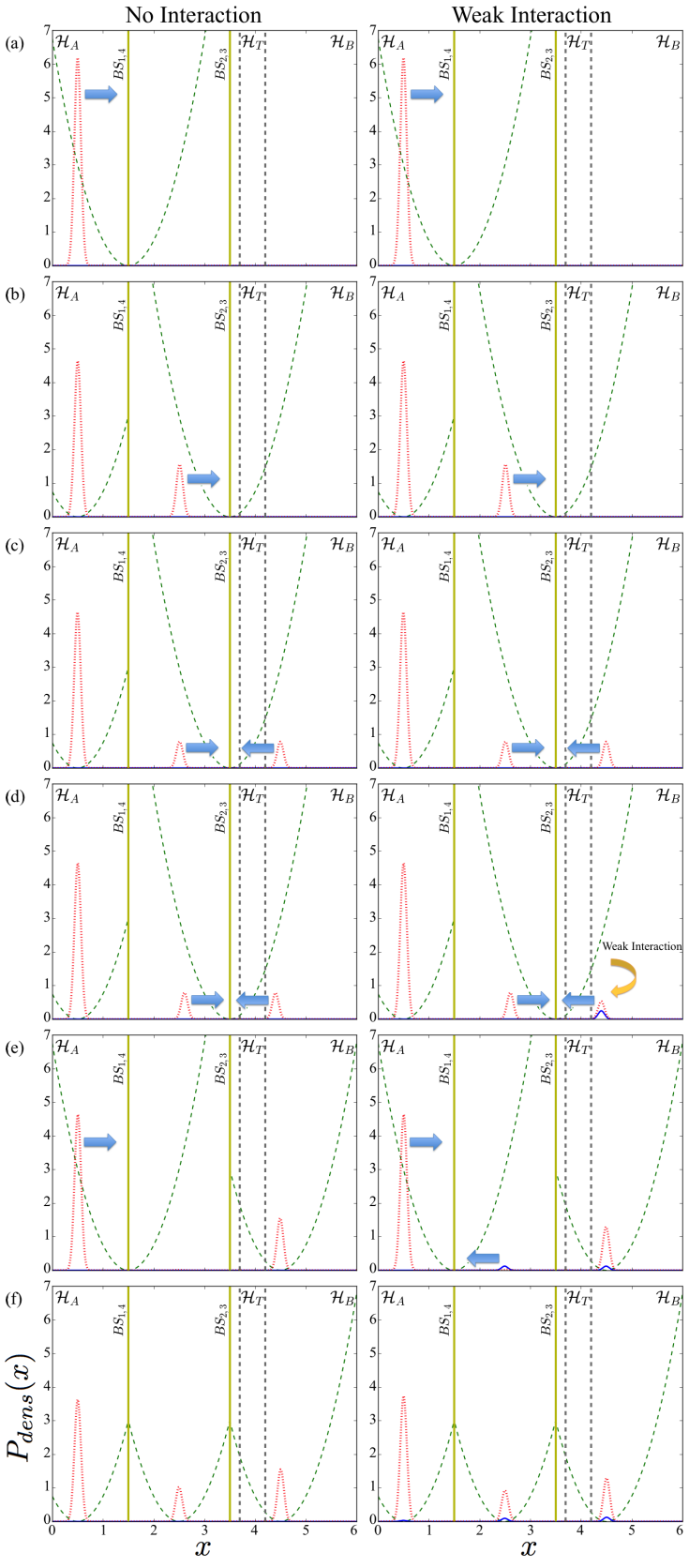


Figure 3.15 The quantum evolution [time steps (a) to (f)] of the probability density distribution of a spin-$\frac{1}{2}$ particle propagating through an NMZI with (right) and without (left) a weak spin-rotation interaction (exaggerated for visibility) in Bob's laboratory. The dotted red and solid blue curves indicate spin up and spin down components of the wavefunction, respectively. The dashed green curves show the potentials. Beamsplitters are denoted with vertical yellow lines. The spatial laboratories are separated by vertical dashed grey lines. This figure originally appeared in Ref. [47].

wavefunction from Figure 3.15(b) has been split by BS_2. This allows a transmitted component to enter Bob's laboratory. In Fig. 3.15(d), the middle and right part of the wavefunction are again about to interact on the right beam-splitter (now BS_3). This is where we introduce the tagging mechanism, mimicking a disturbance in Bob's laboratory by imposing a local magnetic field that rotates the spin of Bob's wavefunction component. This is shown in the right frame, whilst the left frame shows the scenario of perfect quantum channels. Also note that the weak interaction, in this proof-of-principle simulation, has been amplified for visibility purposes. Figure 3.15(e) shows the probability density distribution after the third beam-splitter interaction. It shows that if a weak disturbance takes place in Bob's laboratory (right frame), this will result in the "leakage" of a small part of the wavefunction, from Bob's laboratory back into Alice's. Crucially, this will be true even if we post-select such that we ignore outcomes where the particle is found in Bob's laboratory. After the third beam-splitter interaction, the external potentials are again shifted such that the left and middle part of the wavefunction interact on the left beam-splitter (BS_4), whilst the right part is held in Bob's laboratory. In Fig. 3.15(f) we see the output probability density distribution of the NMZI experiment.

In conclusion, the result of the weak tagging interaction in Bob's laboratory [right frame of Fig. 3.15(d)], is to distort the interference on BS_3. This is visible between Fig. 3.15(d) and Fig. 3.15(f). In the scenario on the left, without any weak interaction in Bob's laboratory, the effect of BS_3 is the constructive interference of the right and the middle wavefunction components, such that they propagate to Bob's laboratory, never to return to Alice's. However, if a weak interaction takes place in Bob's laboratory, this will cause a fraction of the wavefunction [blue curve in the right frame of Fig. 3.15(e)] to propagate back to Alice. This small part of the wavefunction will then interfere with the left part of the wavefunction on BS_4. The result of this is that the output statistics in Alice's laboratory (at $x = 0.5$ and $x = 2.5$) are different between the two scenarios (with and without a tagging interaction). In the scenario with a tagging interaction, the output distribution is dependent on the strength of the tagging. Hence, Alice obtains Fisher information of Bob's tagging interaction. The consequence of this is ultimately that type-I counterfactuality is satisfied only if the NMZI device is constructed from absolutely pure quantum channels.

7.2 A concatenated NMZI

In the section above we investigated the counterfactual violations of the reduced type I scheme. Unfortunately, an extension of the analysis to the full CNMZI scheme, as presented by Salih *et al.* [39], is not straightforward. Previous attempts at evaluating the counterfactuality of this scheme have been subject to heavy criticism (see Refs. [36–45, 79]). This is partially because

of the analytical complication arising from the fact that there are $(N-1)\times(M-1)$ paths in and out of Bob's laboratory (see Fig. 3.11). This complicates the notion of particle *presence* in terms of the analysis in Section 4. Nevertheless, it is possible to calculate the leakage of information from Bob to Alice, with regard to the weak disturbance in Bob's laboratory. In general, this Fisher information can take much larger values than the single-entry scenarios calculated with Eq. 3.21.

Even for smaller values of N and M, the analytical expression for the classical Fisher information takes a complicated and non-informative form. For this reason we proceed with a numerical finite-difference method and obtain a comprehensive approximation of the classical Fisher information flowing from Bob to Alice in the CNMZI setup. This quantity can then be used in Eq. 3.76 in order to calculate the counterfactual violation strength.

Following the reasoning in this chapter, a realistic model of the type I CNMZI device has to account for some disturbances or systematic errors—however minute—occurring within Bob's laboratory. Hence, we change our optical setup of the type I logical 0 process (no D_B) to include tagging interactions within Bob's laboratory (see Fig. 3.16). In Fig. 3.16 weak polarization rotations have been inserted in every path through Bob's laboratory.

Alice will not receive the total Fisher information available from the detection statistics of all the detectors, but a spatially conditioned Fisher information, as described by Eq. 3.78. Figure 3.17 shows the spatially conditioned (i.e. $F \to F_{\boldsymbol{A}}$ in Eq. 3.76) type I counterfactual violation strength, $\mathscr{D}_{\boldsymbol{A}}$, as a function of N and M, under the assumption that Bob does not insert his detectors, and that the polarization rotations are carried out with a weak polarization parameter: $\theta_w = 10^{-6} \ll M^{-1}$. It is easy to see that $\mathscr{D} \geq \mathscr{D}_{\boldsymbol{A}}$, such that Fig. 3.17 represents a lower bound on the type I counterfactual violation strength of a process based on the device of Fig. 3.16.

From Fig. 3.13 it is clear that M and N have to be large in order to implement the type I protocol with low bit error rates. As we can see from Fig. 3.17 such values of N and M generate counterfactual violation strengths that are many orders of magnitude larger than unity. Based on our measure of type I counterfactuality, it is, therefore, fair to conclude that the CFC scheme of Ref. [39] is, *de facto*, not counterfactual.

8 Evaluation of Type II Counterfactuality

We have seen that realistic quantum setups struggle and fail to satisfy type I counterfactuality. Previous indications of this—together with an ambition to simplify the experimental apparatus—were part of my motivation for developing the type II protocol.

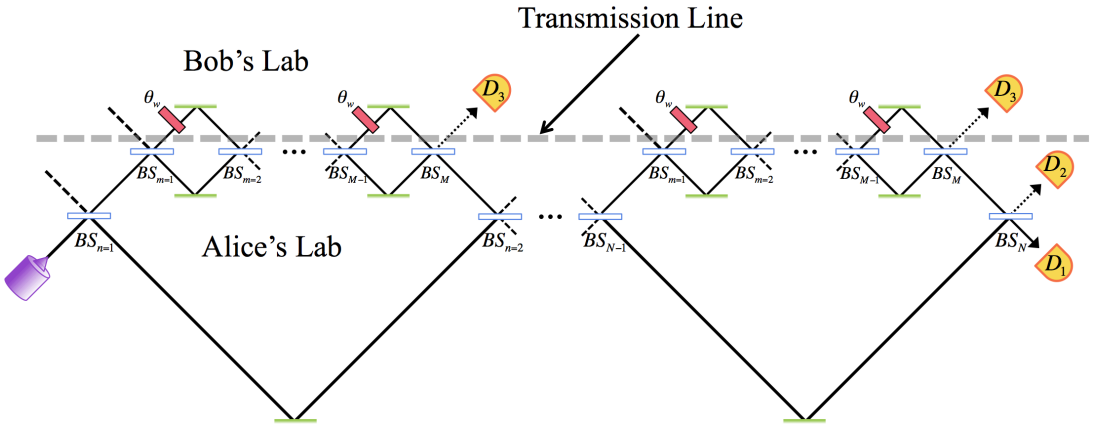


Figure 3.16 The device setup for the logical 0 process for the CNMZI (see Fig. 3.11). Now, tagging interactions (as Eq. 3.9) have been inserted within Bob's laboratory. These weak polarization interactions mimic systematic errors in the quantum channels.

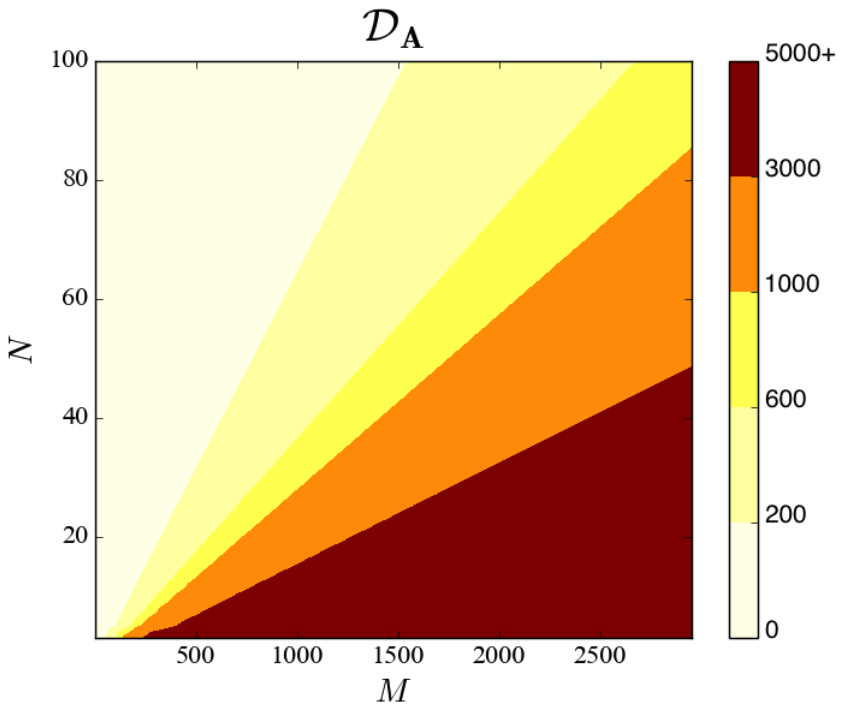


Figure 3.17 The spatially conditioned type I counterfactual violation strength as a function of the beam-splitter numbers N and M in the scenario where Bob does not introduce his detectors in Fig. 3.11. This figure originally appeared in Ref. [47].

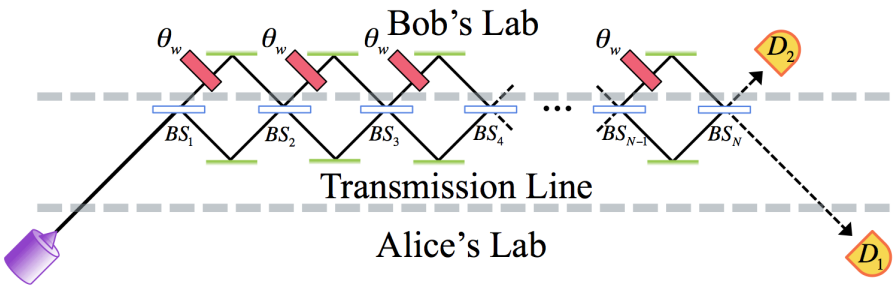


Figure 3.18 Realistic systematic errors in the type II logical 0 process are mimicked by inserting weak polarization rotations within Bob's part of the CMZI.

As described in Chapter 2, the type I and type II protocols' processes are both initiated by Alice inputting a photon into the relevant communication device. In the type II protocol this is a CMZI device (see Fig. 3.12). In the 0-bit processes Bob leaves his laboratory free from absorbing objects. The interrogating particle should then evolve via the CMZI (shared between the transmission line and Bob's laboratory) and eventually be output to Bob's detector D_2, such that Alice never detects it. As type II counterfactuality requires that particles which interact with Bob never propagate to Alice, it is essential that the CMZI evolution is perfect. However, whilst the type I definition of counterfactuality forbids the evolution of interrogating particles from Alice's laboratory to Bob's, this is allowed by the type II definition.

If no specific bit-encoding scheme is used in the type II protocol, the main source of bit errors come from the 1-bit processes.[5] However, whilst the type I CFC scheme requires a huge number, N, of beam-splitters for low bit error rates, a number of $N \approx 100$ generates very good rates in the type II scheme. For bit error rates to be kept below 5 % in the type II 1-bit process, it is sufficient that $N = 50$.

To mimic weak disturbances we again insert weak polarization rotators (representing the tagging mechanism) in the paths that pass through Bob's laboratory (see Fig. 3.18). The counterfactual violation strength of a type II protocol ($P_{\boldsymbol{A}}$ from Eq. 3.77) can then be simulated by a finite difference method.

Two general remarks should be made with regard to $P_{\boldsymbol{A}}$. First, $P_{\boldsymbol{A}}$ grows with increasing numbers of N. Second, $P_{\boldsymbol{A}}$ is dependent on θ_w. Figure 3.19 shows the type II counterfactual violation strength as a function of N for different values of $\theta_w \ll N^{-1}$. From Fig. 3.19 it is obvious that type II counterfactuality is satisfied for small values of θ_w. More specifically, for the values of θ_w explored in the figure, $P_{\boldsymbol{A}}$ is kept well below the free-space interaction value of 1. Nevertheless, it is also evident—and very interesting—that large amounts of classical

[5] For a specific bit-encoding scheme see Chapter 4.5, which allows for the probability of bit error rates to approach zero, even for imperfect beam-splitters and small values of N.

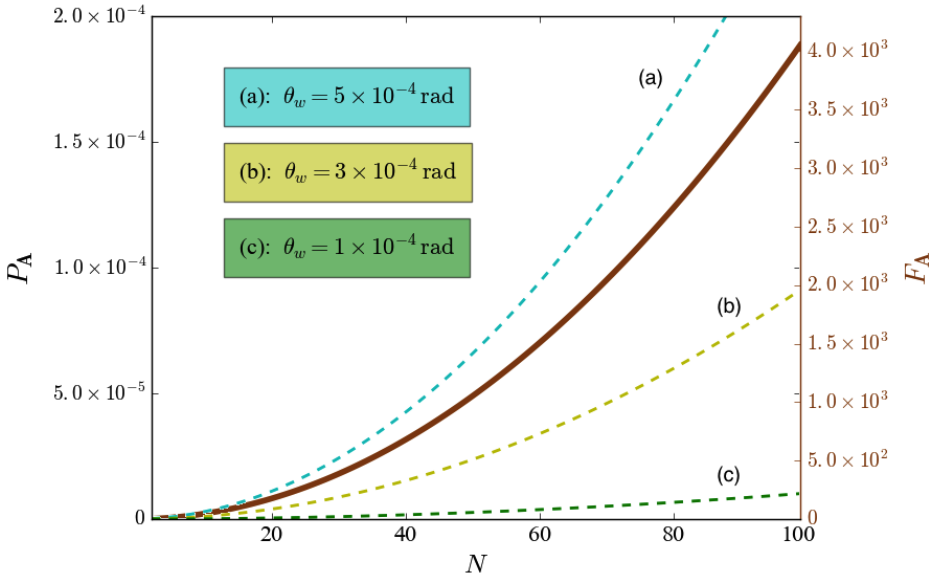


Figure 3.19 The type II logical 0 counterfactual violation strength, $P_{\boldsymbol{A}}$ (dashed lines read on the left vertical axis), and the spatially conditioned Fisher information, $F_{\boldsymbol{A}}$ (solid line read on the right vertical axis), as functions of the beam-splitter number, N.

Fisher information, $F_{\boldsymbol{A}}$, propagate from Bob to Alice, regardless of the small values of $P_{\boldsymbol{A}}$. Note that the value of $F_{\boldsymbol{A}}$ is independent of θ_w (for small θ_w). The value of $P_{\boldsymbol{A}}$ is not.

We can thus conclude this section by stating that the type II protocol, outlined in Chapter 2, satisfies its definition of counterfactuality well—even in the presence of disturbing tagging interactions within Bob's laboratory. For weak tagging interactions, the probability of having a particle propagating from Bob to Alice is close to zero. Interestingly, Alice still receives large amounts of Fisher information of Bob's tagging parameter θ_w.

9 Conclusion

In this chapter we have deployed arguments based on the concept of quantum particle *presence* in order to scrutinise the two CFC protocols outlined in Chapter 2.

Owing to the special role that observations and measurements have in the standard interpretations of quantum mechanics, there is an inherent problem with the definition of inter-measurement presence. We addressed this problem by arguing that absolutely perfect and pure quantum evolutions are impossible, and that any realistic consideration of the interferometers studied in this thesis has to include unwanted disturbances to the wavepacket. Such disturbances will obviously impact the wavefunction propagation through the areas where they act. As the circuits we are studying in this work are non-polarizing, the unwanted

disturbances, named "tagging" interactions, were modelled by weak polarization rotations of the quantum states. We then used the measure of classical Fisher information to arrive at an interesting conclusion: the information of the tagging interaction available at the output detectors is directly proportional to the integrated probability density distribution passing through the location of the tagging. Hence, we can use the classical Fisher information (of certain circuits) to discuss something that, from an information-theoretical position, represents an inter-measurement presence.

Motivated by the increasing interest in the study of NMZI, I have provided a thorough information theoretical study of parameters estimation with these structures. As postulated, it was found that the Fisher information of an unknown interaction, placed in an NMZI, is proportional to the integrated probability density that has passed through the interaction in the Schrödinger picture. This proportionality is, however, not satisfied by the Shannon mutual information if the interaction takes place within the inner MZI.

Given that counterfactual schemes make claims with regards to the presence of particles during the course of the protocol, it is impossible to discuss them from a purely Copenhagen-like perspective. However, the condition for type I counterfactuality to be satisfied—that no particle is allowed to travel from Alice to Bob—can be effectively monitored by the Fisher information measure. As the original type I scheme is based on a complex CNMZI, an analytical study is more informative if carried out on a reduced post-selected scheme based on a single nested MZI. My analysis shows that it is impossible for the reduced scheme to satisfy type I counterfactuality. Moreover, the numerical investigation of the full original protocol results in similar conclusions regarding counterfactual violations. Hence, the beautiful idea of type I counterfactuality seems to be unachievable in any realistic quantum experiment.

Type II counterfactuality, on the other hand, imposes less stringent criteria on the intermediate evolution of the quantum state. The requirement that no wavefunction propagates from Bob to Alice can be tested for by a simple probability measure. A numerical simulation of the type II protocol confirms the possibility of experimentally implementing the protocol with small counterfactual violation rates. As suggested in the previous chapter, the type II protocol is highly robust against disturbances. Even with the introduction of disturbing tagging interactions in Bob's laboratory, counterfactuality is retained.

Chapter 4

Quantum Fisher Information and Particle Presence

1 Chapter Summary

In this chapter I shall extend and generalise my analysis of inter-measurement quantum particle presence. The material is mostly based on preliminary work, which I intend to submit for publication later this year. In Section 3 I shall discuss briefly how one can define a classical object, and how that definition can be used to interpret a quantum object. In Section 4 I provide a theoretical derivation of the quantum Fisher information of an arbitrary two-level interaction in a general quantum circuit. Finally, in Section 5, I use this theory to extend the methodology in Chapter 3, and provide a more general framework for the interpretation of inter-measurement quantum particles. This section also contains an in-depth study of the NMZI from a quantum Fisher information perspective.

2 Setting the Scene: Final Act

From a general point of view, a quantum circuit transforms some input state into an output state. In most scenarios it is the output state that is of interest. However, in Chapter 2, we saw that the intermediate quantum state is highly interesting in the study of counterfactual communication (CFC). This is because the concept of counterfactuality makes ontological claims with respect to the inter-measurement quantum particles. In order to truly understand what counterfactuality or interaction-freeness means, we need to first establish what a quantum object (or particle) really is. Despite the lack of a general consensus on this, quantum physicists have been discussing counterfactual phenomena for decades. This might

seem odd, as a counterfactual process is defined with respect to the paths of the particles that are employed in it. However, the discussion has often centred around the evolution of the quantum wavefunction, which can be mapped out deterministically. But from an ontological perspective, this is not satisfactory; there is no straightforward way to interpret how the mathematical tool of the wavefunction corresponds to the underlying physical reality.

In Chapter 1.5 and Chapter 2.3.2 we saw that it is difficult to use the standard wavefunction picture to interpret the physical reality of a quantum particle in certain post-selected NMZI experiments. Whilst only an infinitesimal component of the wavefunction from the inner MZI propagates to the measurement output, a finite weak value with respect to an interaction in the inner MZI is measurable at the output. Such finite weak traces have been said to describe inter-measurement quantum particles in the TSVF [36].

In Chapter 3 we took an alternative approach to discussing inter-measurement quantum particles. By studying the information which is available at the measurement output of certain optical interferometers we saw that it is possible to learn about where, and to what extent, a quantum particle had been encoded by a tagging interaction. This enabled us to see whether or not a quantum particle had been in Bob's laboratory, and ultimately if the process had defied counterfactuality. In this chapter I shall extend these methods and build a more general theory for the study of inter-measurement quantum presence. Instead of looking at non-polarizing optical interferometers, I shall consider general quantum circuits; and instead of only looking at polarizing tagging mechanisms, I shall allow for any two-level disturbance to represent the tagging. Moreover, I will show that, given weak tagging interactions (which are the premise of a functioning quantum circuit), the theory can be applied to the analysis of several taggings in the same quantum circuit. I will then use this extended framework to revisit the NMZI experiment, and to discuss how the inter-measurement quantum particles can be interpreted from an information-theoretical perspective.

3 A Physical Object

In order to develop a rigid framework for the analysis of inter-measurement quantum presence, we need to have a clear ontological idea of what a quantum object (or particle) is. It is natural to start with a classical discussion before moving on to the quantum scenario, as our predetermined conception of objects and particles relies on our classical observations of the world around us and the corresponding terminology. In this chapter we shall use the word "classical" to describe the deterministic theories of pre-quantum physics—the theories that are remarkably successful in modelling macroscopic systems. In the classical world there is a one-to-one correspondence between experimental input and output, and all processes are

(in principle) reversible. Moreover, there is a direct equivalence between theoretical states described in models, the states revealed by measurements, and the underlying *real* physical state. The same is not necessarily true in quantum physics. For example, in the standard Copenhagen interpretation of quantum mechanics, measurements impose a non-unitary evolution of the wavefunction, which makes the theory irreversible.

3.1 A classical object

It might seem simple to answer the question: What is a classical object? However, the notion of classical objects is so general, that it is difficult to arrive at a definition that is physically useful. In order to keep in line with previous discussions in this thesis, I shall be approaching the question from an operational and information-theoretical perspective.

With regard to our knowledge of the world, the three dimensions of space are different to all other degrees of freedom—all processes, which directly or indirectly, provide us with information are essentially measurements of a given region of space [20].

From a classical point of view, and at the most elementary level, a volume of space can exist in one out of two states: occupied or empty. These two states are fundamentally different. Whilst the state of occupation allows for the investigation of a much richer set of internal states (such as colour, temperature or mass), the state of emptiness only tells us that there is no more information to gain from the volume of interest. Whilst the state of occupation can undergo internal changes, the state of emptiness cannot.

For the purposes of this thesis, I argue that a useful way to define a classical object is as a region of three-dimensional space that is *occupied*. Crucially, it is the object's capability to provide additional *internal* information which distinguishes it from emptiness.

Classical objects are, by definition, large enough for quantum interferences to be negligible—being compositions of large numbers of smaller elementary particles. Hence, any observation of a classical object is by default an interrogation of many particles. Independently of one's interpretation of quantum mechanics, the lack of quantum interference of large classical objects ensures the appeared determinism in observations of them. In experimental scenarios, where we indeed do experience apparent indeterminism, it is either because the systems are too complicated to make predictions about, or simply because the observed objects are quantum.

3.2 A quantum object

Previously in this thesis we discussed quantum particles with reference to their respective measurement probability distributions. Experimentally, knowledge of the outcome probability

distributions requires the evaluation of a large number of identically prepared quantum systems. Whilst a classical object allows us to determine its physical properties exactly by direct measurements, the determination of the physical properties of a quantum particle requires the repeated measurement of the same properties. This is best understood in the Schrödinger picture, in which a quantum state can evolve into a superposition of observable eigenstates. The 4^{th} postulate ensures that a single measurement only yields the probabilistic observation of one of these eigenstates.

One can consider the repeated measurement process of quantum particles, and the single one of classical objects, to be rather similar, as the measurement of a classical object implicitly encompasses the measurement of a large number of particles. I would argue that the crucial difference between our perception of quantum and classical objects lies in the possibility of seemingly contradictory results of repeated measurements on identical quantum systems.

The measurement outcomes of identically prepared classical objects always yields information with respect to the same volumes of space; but the measurement outcomes of identically prepared quantum objects (particles) can yield information with respect to different volumes of space. Consider a photon incident on the lower input path of a beam-splitter and subsequently detected at one of the two output paths [see Fig. 1.10(a)]. Depending on which detector is triggered, the experimentalist can only gain information (about, for example, a polarization interaction) from the corresponding output path. How can we then use the classical concept of an object to describe something quantum? A possible solution is to accept the suggestion above, that the need for repeated measurements to acquire the information of a quantum particle is not too dissimilar from the measurement of a classical object. The spatial extent of a quantum object can then be thought of as the volume of space which, after repeated measurements on identical systems, provides information regarding the properties of the physical state which occupies this region.[1] By doing this an object becomes defined by the information we can obtain about it rather than by individual measurements of it.

Finally, we need to address the question of inter-measurement presence of an object. Classically, there is no problem associated with the postulation of the inter-measurement path of an object, as all trajectories are deterministic. However, as we have seen, the strong measurement of a quantum system alters its future interference. The information approach to the study of objects solves this issue. As we saw in Chapter 3, the information with respect to an interaction can be measured at the output of an experiment. Hence, if one knows the

[1] Note that quantum wavefunctions and probability density distributions, in general, extend to infinity in the continuum limit. A thorough discussion regarding the vagueness or indeterminacy of quantum objects can be found in Refs. [20, 110, 111]. However, in this thesis we are interested in discrete quantum states, for which we can consider the spatial extent of wavefunctions to be finite and well-defined.

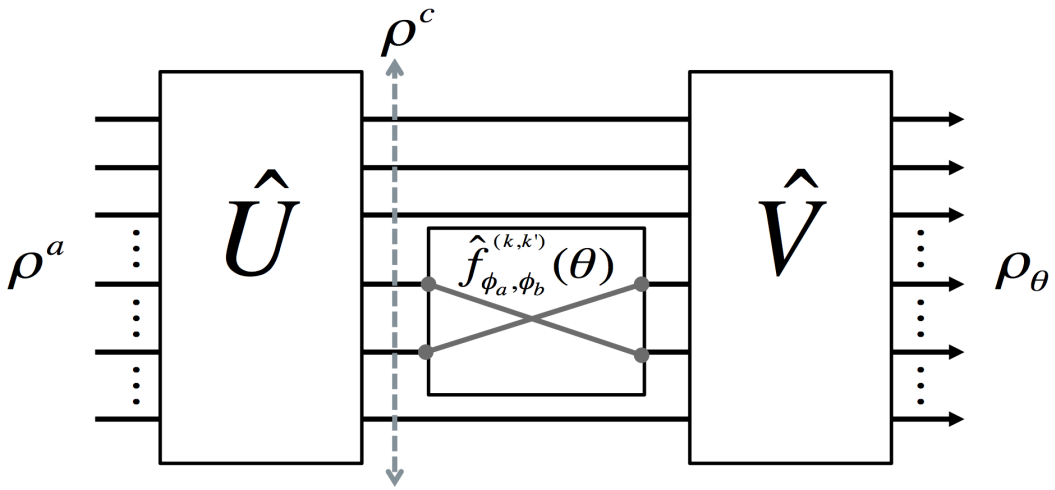


Figure 4.1 Sketch of a general discrete quantum circuit of the form of Eq. 4.1, which is described in the main text.

spatial location of the various interactions, one can use the output information to map out the quantum particle's past presence at these locations. In the following sections of this chapter I shall extend the use of the tagging mechanisms from Chapter 3 to explore the possibilities of using the information available at the experiment output to map out the physical paths of quantum particles. We shall see that the methodology that was fruitful in the analysis of the CFC schemes can be extended to general discrete quantum circuits.

4 Quantum Fisher Information - General Circuits

In this section I shall further generalise the study of the propagation of an n-level quantum system through a two-level interaction. In Chapter 3.3.3 we saw how the quantum Fisher information provides a measure of the maximum amount of information that can be obtained of the parameter θ from a general quantum state ρ_θ, given that the most optimal measurement is conducted. The information encoding setup (state preparation) of the initial state ρ^a is shown in Fig. 4.1. However, in contrast to the analysis of Chapter 3 (and Appendix B), we require that the quantum state is not collapsed directly after $\hat{V}$, but that the most information optimal measurement is conducted. The question we are interested in answering is: How well can we estimate the variable θ that sets a single two-level interaction in an n-level quantum circuit?

We assume that an input state, ρ^a (now a density matrix) evolves through the quantum circuit depicted in Fig. 4.1 and is output as ρ_θ. The two-level interaction is represented by

the unitary $\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)$. In order to calculate the quantum Fisher information available from the output state we need a number of definitions.

Definitions:

1. Let $\rho^a = |\psi^a\rangle\langle\psi^a|$ be density matrix of the pure quantum state which is the input of the experiment.

2. Let the quantum circuit be divided into three steps such that the unitary operation of the circuit is $\hat{V}\cdot\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)\cdot\hat{U}$, where $\hat{U}$ and $\hat{V}$ are known; and $\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)$ is the two-level interaction, acting between levels k and k', set by the unknown parameter θ. The evolution of the initial quantum state can then be modelled by

$$\mathscr{S}\rho^a\mathscr{S}^\dagger \equiv \left[\hat{V}\cdot\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)\cdot\hat{U}\right]\rho^a\left[\hat{V}\cdot\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)\cdot\hat{U}\right]^\dagger \equiv \rho_\theta. \tag{4.1}$$

3. Let $\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)$ be given by

$$\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & e^{i\phi_a}f_{k,k}(\theta) & \cdots & e^{i\phi_b}f_{k,k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-i\phi_b}f_{k',k}(\theta) & \cdots & e^{-i\phi_a}f_{k',k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \tag{4.2}$$

   where $\phi_{a,b}$ sets the unitary phases, and we require that $f_{k,k}(\theta) = \sqrt{1-\left[f_{k,k'}(\theta)\right]^2}$, $f_{k',k}(\theta) = -f_{k,k'}(\theta)$ and $f_{k',k'}(\theta) = f_{k,k}(\theta)$, with $f_{k,k}(\theta) : \mathbb{R}\to[-1,1]$. In the following I shall use a simplified notation and define: $\hat{f}^{(k,k')}_\theta \equiv \hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)$.

4. Let $\rho^c = |\psi^c\rangle\langle\psi^c| \equiv \hat{U}\rho^a\hat{U}^\dagger$ denote the intermediate quantum state just before the interaction takes place.

5. Let ψ^c_i be the wavefunction component of ρ^c at the i^{th} level.

6. Let $\Phi_Q \equiv \phi_a - \phi_b + \arg(\psi^c_k) - \arg(\psi^c_{k'})$ define a phase which we shall refer to as the "circuit phase".

With these definitions it is possible to derive an expression for the quantum Fisher information of the parameter θ. In the following we use a simplified notation for the derivative with respect to θ: $\dot{x} \equiv \frac{\partial}{\partial\theta}x$.

Theorem 1. *The quantum Fisher information of the two-level interaction parameter, θ, encoded in the n-level quantum state, ρ_θ, via a general quantum circuit, is given by*

$$F_Q(\theta) = \left[4\left(|\psi_k^c|^2 + |\psi_{k'}^c|^2\right) - 16|\psi_k^c|^2|\psi_{k'}^c|^2\sin^2(\Phi_Q)\right]\left(\frac{\dot{f}_{k,k'}}{f_{k,k}}\right)^2. \tag{4.3}$$

Proof. From Eq. 3.4 we know that the quantum Fisher information is given by

$$F_Q(\theta) = \mathrm{Tr}[\rho_\theta \Lambda_\theta^2], \tag{4.4}$$

where Λ_θ is the logarithmic derivative operator, implicitly defined by

$$\frac{\partial}{\partial\theta}\rho_\theta = \frac{1}{2}(\Lambda_\theta\rho_\theta + \rho_\theta\Lambda_\theta). \tag{4.5}$$

We can express our output state as $\rho_\theta = \hat{V}\hat{f}_\theta^{(k,k')}\hat{U}\rho^a\hat{U}^\dagger\hat{f}_\theta^{(k,k')\dagger}\hat{V}^\dagger$ or, using Definition 4, $\rho_\theta = \hat{V}\hat{f}_\theta^{(k,k')}\rho^c\hat{f}_\theta^{(k,k')\dagger}\hat{V}^\dagger$. Since ρ_θ is a pure state, we know that:

$$\frac{\partial}{\partial\theta}\rho_\theta = \frac{\partial}{\partial\theta}\rho_\theta^2 = \dot{\rho}_\theta\rho_\theta + \rho_\theta\dot{\rho}_\theta, \tag{4.6}$$

such that $\Lambda_\theta = 2\dot{\rho}_\theta$. The quantum Fisher information is thus given by

$$F_Q(\theta) = 4\mathrm{Tr}[\rho_\theta\dot{\rho}_\theta^2]. \tag{4.7}$$

We temporarily drop the (k,k')-superscripts and express the logarithmic derivative in terms of the unitary operations of the quantum circuit:

$$\Lambda_\theta = 2\dot{\rho}_\theta = 2(\hat{V}\dot{\hat{f}}_\theta\hat{U}\rho^a\hat{U}^\dagger\hat{f}_\theta^\dagger\hat{V}^\dagger + \hat{V}\hat{f}_\theta\hat{U}\rho^a\hat{U}^\dagger\dot{\hat{f}}_\theta^\dagger\hat{V}^\dagger). \tag{4.8}$$

We can then express the Fisher information in terms of the unitary operations and the input and output states:

$$F_Q(\theta) = 4\mathrm{Tr}[\rho_\theta(\hat{V}\dot{\hat{f}}_\theta\hat{U}\rho^a\hat{U}^\dagger\hat{f}_\theta^\dagger\hat{V}^\dagger + \hat{V}\hat{f}_\theta\hat{U}\rho^a\hat{U}^\dagger\dot{\hat{f}}_\theta^\dagger\hat{V}^\dagger)^2], \tag{4.9}$$

which after some algebra and the use of Definition 4, simplifies to

$$\begin{aligned} F_Q(\theta) =& 4\mathrm{Tr}[(\rho^c \hat{f}_\theta^\dagger \dot{\hat{f}}_\theta)^2] + 4\mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \hat{f}_\theta)^2] \\ &+ 4\mathrm{Tr}[(\rho^c \hat{f}_\theta^\dagger \dot{\hat{f}}_\theta)(\rho^c \dot{\hat{f}}_\theta^\dagger \hat{f}_\theta)] + 4\mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \dot{\hat{f}}_\theta)]. \end{aligned} \tag{4.10}$$

The operator $(\dot{\hat{f}}_\theta^\dagger \hat{f}_\theta)$ is skew-Hermitian, so we can simplify our expression further:

$$F_Q(\theta) = 4\mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \hat{f}_\theta)^2] + 4\mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \dot{\hat{f}}_\theta)]. \tag{4.11}$$

The first trace in Eq. 4.11 can be evaluated by explicitly calculating $\dot{\hat{f}}_\theta^\dagger \hat{f}_\theta$, which contains only four entries. The trace can then be expressed in terms of the components of the mid-circuit quantum state, $\rho^c = |\psi^c\rangle \langle \psi^c|$, such that

$$\mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \hat{f}_\theta)^2] = -4|\psi_k^c|^2 |\psi_{k'}^c|^2 \sin^2(\Phi_Q) \left(\frac{\dot{f}_{k,k'}}{f_{k,k}} \right)^2, \tag{4.12}$$

where we use Definition 6:

$$\Phi_Q \equiv \phi_a - \phi_b + \arg(\psi_k^c) - \arg(\psi_{k'}^c). \tag{4.13}$$

The second trace in Eq. 4.11 is given by

$$\begin{aligned} \mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \dot{\hat{f}}_\theta)] =& \rho_{k,k}^c (\dot{f}_{k,k}^* \dot{f}_{k,k} + \dot{f}_{k',k}^* \dot{f}_{k',k}) + \rho_{k',k'}^c (\dot{f}_{k,k'}^* \dot{f}_{k,k'} + \dot{f}_{k',k'}^* \dot{f}_{k',k'}) \\ &+ \rho_{k,k'}^c (\dot{f}_{k,k'}^* \dot{f}_{k,k} + \dot{f}_{k',k'}^* \dot{f}_{k',k}) + \rho_{k',k}^c (\dot{f}_{k,k}^* \dot{f}_{k,k'} + \dot{f}_{k',k}^* \dot{f}_{k',k'}). \end{aligned} \tag{4.14}$$

As ρ^c is pure and $\hat{f}_\theta$ is unitary, it follows that:

$$\begin{aligned} \mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \dot{\hat{f}}_\theta)] =& \rho_{k,k}^c (\dot{f}_{k,k}^* \dot{f}_{k,k} + \dot{f}_{k',k}^* \dot{f}_{k',k}) + \rho_{k',k'}^c (\dot{f}_{k,k'}^* \dot{f}_{k,k'} + \dot{f}_{k',k'}^* \dot{f}_{k',k'}) \\ &+ \rho_{k,k'}^c (\dot{f}_{k,k'}^* \dot{f}_{k,k} + \dot{f}_{k',k'}^* \dot{f}_{k',k} + \dot{f}_{k,k}^* \dot{f}_{k,k'} + \dot{f}_{k',k}^* \dot{f}_{k',k'}) \\ =& \rho_{k,k}^c (\dot{f}_{k,k}^* \dot{f}_{k,k} + \dot{f}_{k',k}^* \dot{f}_{k',k}) + \rho_{k',k'}^c (\dot{f}_{k,k'}^* \dot{f}_{k,k'} + \dot{f}_{k',k'}^* \dot{f}_{k',k'}) \\ =& (\rho_{k,k}^c + \rho_{k',k'}^c)(\dot{f}_{k,k}^* \dot{f}_{k,k} + \dot{f}_{k',k}^* \dot{f}_{k',k}). \end{aligned} \tag{4.15}$$

By using the relations of the elements of $\hat{f}_\theta$, we find that

$$\mathrm{Tr}[(\rho^c \dot{\hat{f}}_\theta^\dagger \dot{\hat{f}}_\theta)] = (|\psi_k^c|^2 + |\psi_{k'}^c|^2) \left(\frac{\dot{f}_{k,k'}}{f_{k,k}} \right)^2. \tag{4.16}$$

By combining Eq. 4.12 and 4.16 we obtain the quantum Fisher information in the form of Eq. 4.3 as given in Theorem 1:

$$F_Q(\theta) = \left[4\left(|\psi_k^c|^2 + |\psi_{k'}^c|^2\right) - 16|\psi_k^c|^2|\psi_{k'}^c|^2 \sin^2(\Phi_Q)\right]\left(\frac{\dot{f}_{k,k'}}{f_{k,k}}\right)^2.$$

□

Theorem 1 shows a beautiful and simple relation between the wavefunction and the quantum Fisher information. It turns out that the information encoded in a quantum particle from a two-level interaction is always a function of the two probability density components between which the interaction acts. Earlier in this thesis, we considered single particle wavefunctions, but the result of Eq. 4.3 holds for any pure quantum state acted upon by a two-level interaction. This includes, for example, several sets of universal quantum gates.[2] It is remarkable that the complicated and convoluted formula for the quantum Fisher information can be simplified to Eq. 4.3 for some of the most valuable operations of quantum computers. An example of an experimental application of this is for calculating how well a given quantum computer can estimate the fidelity of an additional two-qubit quantum gate.

The first term of Eq. 4.3 is directly proportional to the probability of finding the particle in any of the two interaction levels if detectors had been placed just where the interaction takes place. This term connects (the modulus square of) the wavefunction representation of quantum particles to the information content obtained via optimised quantum experiments. The second term comes across as less appealing. It can reduce the Fisher information by a quantity set by the circuit phase. In Chapter 3 we saw how the Fisher information measure of interferometers where the second term of Eq. 4.3 is zero extends our ability to discuss the inter-measurement presence of quantum states. But how does a non-zero second term alter this ability? Let us analyse the physical implication and meaning of it.

The circuit phase

The first thing to note with regard to the circuit phase, is that the multiplicative term of $|\psi_k^c|^2|\psi_{k'}^c|^2$ ensures that it can only affect the Fisher information encoding if both the quantum levels involved in the tagging interaction have non-zero occupation prior to the tagging. The more equally the two levels are occupied, the more destructive the impact of the circuit phase can be on the information encoding capacity of the experiment. Note that scenarios where

[2] The entries of $\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta)$ have to be altered slightly depending on which quantum gate one wishes to study. However, the same methodology still holds.

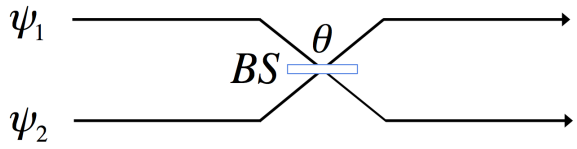


Figure 4.2 Sketch of a simple quantum circuit consisting of only two levels that interact via a beam-splitter set by the unknown parameter θ.

both tagging levels are occupied exclude the ones considered in Chapter 3, where we dealt with polarization tagging of non-polarizing circuits.

Let us, for the sake of investigation, consider scenarios where both the tagging levels are occupied. The simple quantum circuit depicted in Fig. 4.2 can help us in understanding the importance of the second term of Eq. 4.3.

Figure 4.2 shows the simplest of quantum circuits. It consists of only two quantum levels. The unitaries are set such that $\hat{U} = \hat{V} = \hat{1}$, and the tagging mechanism is given by a beam-splitter operation:

$$\hat{f}_{\theta}^{(1,2)} = \begin{pmatrix} e^{i\phi_a}\cos(\theta) & e^{i\phi_b}\sin(\theta) \\ -e^{-i\phi_b}\sin(\theta) & e^{-i\phi_a}\cos(\theta) \end{pmatrix}. \tag{4.17}$$

The input state, ρ^a, will be the same as the mid-circuit state, ρ^c, given by $\rho^c = |\psi_c\rangle\langle\psi_c|$, where $|\psi_c\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$. The circuit phase (Eq. 4.13) is given by $\Phi_Q = \phi_a - \phi_b$. Using Eq. 4.3 we can calculate the quantum Fisher information to be

$$F_Q = 4 - 4\sin^2(\phi_a - \phi_b) = 4\cos^2(\phi_a - \phi_b). \tag{4.18}$$

Equation 4.18 highlights the importance of the circuit phase. In order to determine the Fisher information encoding capacity of a quantum circuit it is not enough to consider just the location of the tagging mechanism. If we, for the sake of argument, set the phases in Eq. 4.18 to $\phi_a = 0$ and $\phi_b = \pi/2$, the tagging interaction is given by

$$\hat{f}_{\theta}^{(1,2)} = \begin{pmatrix} \cos(\theta) & i\sin(\theta) \\ i\sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{4.19}$$

Then, the output state after the interaction is given by $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, where $|\psi_\theta\rangle = \frac{1}{\sqrt{2}}[\cos(\theta) + i\sin(\theta)](|1\rangle + |2\rangle) = \frac{1}{\sqrt{2}}e^{i\theta}(|1\rangle + |2\rangle)$. This makes it clear why the complex phases play such an important role in determining the Fisher information. In some scenarios the circuit phase can result in the "encoding" of the unknown parameter, θ, occurring in the complex phase of the wavefunction, rather than in the quantum levels of it. As we

saw in Chapter 1, the Copenhagen interpretation suggests that the wavefunction and its complex phases are mathematical tools for the calculation of the physically real measurement outcomes (in the circuit case: observable eigenstates of the quantum levels), rather than elements of physical reality. Using parameter estimation, we can only physically measure phase-differences, and in the scenario considered here, both the $|1\rangle$ and $|2\rangle$ levels share the same phase—that is, $e^{i\theta}$. Hence, there is no relative phase to estimate, and the information of θ is lost—or perhaps it never existed.

One can also consider the circuit phase from an operational point of view. There are two phase difference components that determine the circuit phase. The first phase difference, $\phi_a - \phi_b$, is set by the tagging interaction itself. The second phase difference, $\arg(\psi_k^c) - \arg(\psi_{k'}^c)$, is determined by the wavefunction components just before the tagging; it is indirectly set by the quantum circuit and the input state. The sum of these components influences the Fisher information encoding capacity of the experiments. Indeed, certain combinations lead to a maximum amount of Fisher information being encoded, and other leads to a minimum, irrespective of the value of θ. We note that for any quantum circuit, there exists a tagging mechanism, $\hat{f}_\theta$, which maximises $F_Q(\theta)$ by its choice of complex phases, ϕ_a and ϕ_b. If this is done, the amount of Fisher information encoded in the quantum particle is always proportional to the integrated probability density at the point of the tagging interaction:

$$\max_{\phi_a,\phi_b} F_Q(\theta) = 4\left(|\psi_k^c|^2 + |\psi_{k'}^c|^2\right)\left(\frac{\dot{f}_{k,k'}}{f_{k,k}}\right)^2. \tag{4.20}$$

But how can we interpret the tagging interactions in the scenarios of vanishing Fisher information? From an operational viewpoint an experiment with an information destructive circuit phase simply means that, despite the presence of a component capable of transforming a quantum state, the output probabilities are less affected by it than in an experiment with an information constructive circuit phase. Conclusively, the operational meaning of a quantum operation is dependent on the pre-operation quantum state, ρ^c.

In fact, the quantum Fisher information is closely related to the *quantum speed limit*, a theorem which relates the instantaneous "speed" (θ serving the role of time) at which two neighbouring output states, ρ_θ and $\rho_{\theta+d\theta}$, are separated under the action of the quantum circuit. The square-root of the quantum Fisher information is proportional to the instantaneous speed of separation [101, 112]. In the light of the quantum speed limit, a destructive circuit phase simply reduces this speed.

According to the Schrödinger picture of quantum mechanics, the wavefunction will still pass through the location of the tagging in Fig. 4.2, even if no measurable tagging takes place. However, the fact that no measurable effect is induced by a tagging means

that it did not result in any interference. Moreover, as has been argued by Lewis (see Ref. [20]), interactions are negligible in the absence of interference. If one follows this line of argument, the scenarios where the circuit phase reduces the Fisher information encoding represents tagging interactions of *weaker strength*. This is perhaps the best way to interpret the circuit phase. However, care should be taken when using the word "strength" here. In common quantum language the word "strength" refers to the value of the tagging parameter, θ. However, here the strength is with regard to an experiment's ability to decode information from a quantum state.[3] One should give special notice to the strength-reducing effect of the circuit phase when using the Fisher information measure to discuss inter-measurement particle presence in general quantum circuits.

5 Inter-Measurement Presence Revisited

The ability to conduct self-interference experiments with single particles is an important feature that clearly distinguishes the quantum world from the classical. The Fisher information measure of a quantum particle in such experiments suggests the possibility of particles existing to different extents at separate locations of space. This is clearly different from our perception of classical particles. The way to measure the "extent" of occupation at a given location is by evaluating the Fisher information of a tagging interaction at this location. This tagging interaction must be such that the circuit phase does not act destructively. However, as all degrees of freedom of a quantum particle constantly interact with their surroundings, one can argue that this is bound to be satisfied in a realistic scenario. But what if we are interested in knowing whether or not a particle has had a past at several separated locations? Can our formalism be extended to include multiple taggings at different locations?

For a quantum circuit to be adequately represented by the scattering matrix, $\mathscr{S}$, from Eq. 4.1, it is crucial that the tagging interactions are very weak. We can define the zero point of θ such that $f_{k,k'}(\theta = 0) = 0$ and $f_{k,k}(\theta = 0) = 1$. Note that also ϕ_a and ϕ_b should be close to 0 for a tagging to have a weak impact on the circuit phase. We would then have $\hat{f}^{(k,k')}_{\phi_a=0,\phi_b=0}(\theta \approx 0) \approx \hat{1}$, such that $\mathscr{S} \approx \hat{V} \cdot \hat{U}$. Finally, if we define our unknown tagging rotation such that $\dot{f}_{k,k'}(\theta = 0) = 1$, the expression for the Fisher information simplifies and becomes independent of θ at $\theta \approx 0$:

$$\lim_{\theta \to 0} F_Q(\theta) = 4\big(|\psi^c_k|^2 + |\psi^c_{k'}|^2\big) - 16|\psi^c_k|^2|\psi^c_{k'}|^2 \sin^2(\Phi_Q). \tag{4.21}$$

[3] Note that this information decoding strength, in general, can be a function of θ (via the last multiplicative term of Eq. 4.3) in addition to its dependence on the circuit phase.

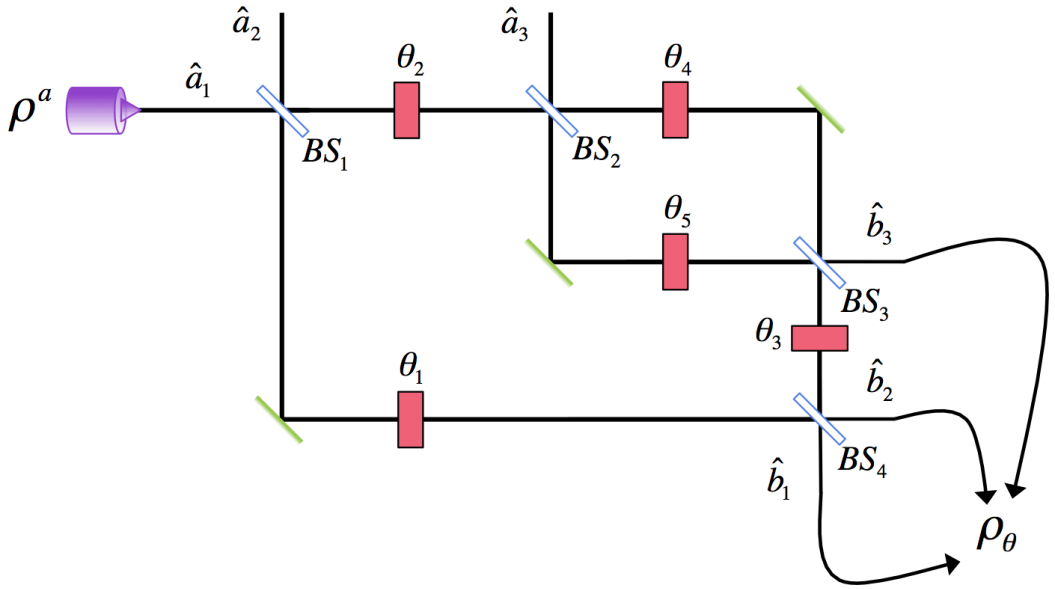


Figure 4.3 Five tagging polarization rotations (satisfying the properties described in the main text) are included at positions, 1-5, in an NMZI. The output state, $\rho_{\boldsymbol{\theta}}$, is measured with an optimal positive-operator valued measure.

In Chapter 3, we saw several examples of where the tagging interaction was vanishingly small, but where the Fisher information was finite. Mathematically, this is caused by the role of the derivative with respect to θ in the Fisher information formula. As the Fisher information with respect to a specific variable only contains a derivative with respect to that variable, the joint consideration of several weak tagging interactions at different locations of the quantum circuit does not affect their respective components of the Fisher information. This allows us to use many weak tagging interactions in order to map out the spatial path of propagation of a quantum particle between measurements. We let θ be transformed to a vector of the parameters of all the tagging interactions: $\theta \rightarrow \boldsymbol{\theta} \equiv (\theta_1, \theta_2, ..., \theta_n)$.

5.1 Quantum particle paths through an NMZI

As in Chapter 3.5, we consider the input state of the NMZI to be $|\psi_{in}\rangle = \hat{a}^\dagger_{1,H}|0\rangle$, such that $\rho^a = |\psi_{in}\rangle\langle\psi_{in}|$. However, this time we consider all the five tagging interactions of Fig. 4.3 to be jointly present in the device. We also let these tagging interactions be set by $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_5)$. As in Chapter 3, we set $r_2 = t_2 = r_3 = t_3 = \frac{1}{\sqrt{2}}$. Our quantum levels are labelled such that θ_1 sets the interaction between $n = 1$ and $n = 4$, $\theta_{2,3,5}$ set interactions between $n = 2$ and $n = 5$, and θ_4 sets the interaction between $n = 3$ and $n = 6$. The output state of the interferometer is given by the joint action of every operation of the circuit on the input state:

$$\begin{aligned}\rho_\theta =&\hat{U}_{BS_4}\hat{f}_{0,0}^{(2,5)}(\theta_3)\hat{U}_{BS_3}\hat{f}_{0,0}^{(3,6)}(\theta_4)\hat{f}_{0,0}^{(2,5)}(\theta_5)\hat{U}_{BS_2}\hat{f}_{0,0}^{(2,5)}(\theta_2)\hat{f}_{0,0}^{(1,4)}(\theta_1)\hat{U}_{BS_1}\cdot\rho^a\\ &\cdot\hat{U}_{BS_1}^\dagger\left[\hat{f}_{0,0}^{(1,4)}(\theta_1)\right]^\dagger\left[\hat{f}_{0,0}^{(2,5)}(\theta_2)\right]^\dagger\hat{U}_{BS_2}^\dagger\left[\hat{f}_{0,0}^{(2,5)}(\theta_5)\right]^\dagger\left[\hat{f}_{0,0}^{(3,6)}(\theta_4)\right]^\dagger\hat{U}_{BS_3}^\dagger\left[\hat{f}_{0,0}^{(2,5)}(\theta_3)\right]^\dagger\hat{U}_{BS_4}^\dagger.\end{aligned} \tag{4.22}$$

According to the generalised analysis of this chapter, the measurements should be optimised with respect to the quantum Fisher information. Fortunately, our measurement of the number of horizontally and vertically polarized photons at each output path represents such an optimised measurement in the NMZI experiment.

Let us start by calculating the quantum Fisher information obtained by experimental evaluation of the device of Fig. 4.3, with respect to the five different unknown parameters. We find that

$$F_Q(\theta_1) = 4r_1^2\frac{1}{1-\theta_1^2}, \tag{4.23}$$

$$F_Q(\theta_2) = 4t_1^2\frac{1}{1-\theta_2^2}, \tag{4.24}$$

$$F_Q(\theta_3) = 2t_1^2\frac{1-\theta_4\theta_5-\sqrt{(1-\theta_4^2)(1-\theta_5^2)}}{1-\theta_3^2}, \tag{4.25}$$

$$F_Q(\theta_4) = 2t_1^2\frac{1}{1-\theta_4^2}, \tag{4.26}$$

$$F_Q(\theta_5) = 2t_1^2\frac{1}{1-\theta_5^2}. \tag{4.27}$$

All of these quantities, except $F_Q(\theta_3)$, correspond exactly to our previous analysis in Chapter 3.5. Hence, all the spatial components of Fig. 4.3 are occupied to the same extent as when the taggings are introduced individually—except for the location of the third tagging. Interpreted with the Schrödinger picture, the reason for the non-zero particle path joining BS_3 and BS_4 (via θ_3) is the destruction of the perfect interference of the inner MZI which takes place if $\theta_4 \neq 0 \neq \theta_5$ and $\theta_4 \neq \theta_5$. Hence, the occupation of this location is not owing to the interferometer itself, but caused by disturbances at the fourth or fifth location. However, for the taggings not to alter the interferometer significantly we require that $\boldsymbol{\theta} \approx 0$. Using Eq. 4.21 the quantum Fisher information of the respective tagging interactions can then be evaluated to be:

$$\lim_{\boldsymbol{\theta}\to 0} F_Q(\theta_1) = 4r_1^2, \tag{4.28}$$

$$\lim_{\boldsymbol{\theta}\to 0} F_Q(\theta_2) = 4t_1^2, \tag{4.29}$$

$$\lim_{\boldsymbol{\theta}\to 0} F_Q(\theta_3) = 0, \tag{4.30}$$

$$\lim_{\boldsymbol{\theta}\to 0} F_Q(\theta_4) = 2t_1^2, \tag{4.31}$$

$$\lim_{\boldsymbol{\theta}\to 0} F_Q(\theta_5) = 2t_1^2. \tag{4.32}$$

We see that all the quantities now correspond with the analysis of Chapter 3.5.

From an ontological perspective the analysis above does not seem implausible; that is, if we are willing to accept that a real quantum object is related to its wavefunction, and simultaneously can exist at different locations to different extents.[4] Even though the quantum particle seems to have occupied several spatial locations at the same point in time, its individual components appear to follow well-defined and continuous paths. So what is it that has caused the NMZI to be subject to such intense scrutiny? The answer lies in the results one obtains when imposing post-selection on the NMZI experiment.

Example: the NMZI revisited

In Chapter 1 we saw that if a quantum particle's past path is defined by its weak trace, a post-selected NMZI experiment will lead to the appearance of discontinuous paths of the particle. This conclusion seems to hinge upon the fringe interpretation of quantum mechanics that is the TSVF (see Chapter 1.5). Let us instead study this phenomenon from our information-theoretical and operational interpretation.

Consider an observer, Alice, who occupies the spatial location around the $\hat{b}_1$ output of Fig. 4.3, and who is not allowed to communicate with the outside world. Alice's colleague, Bob, controls a laser and can input photons to the device. Alice knows that Bob only ever inputs particles in the $|\psi_{in}\rangle$ state, but she does not know when. The only time Alice is certain that Bob sent a photon is when her detector (D_1 from Chapter 3) ticks. An external observer, overlooking the entire experiment would find that

$$\begin{aligned} P(n_1^H = 1|\boldsymbol{\theta}) = &\frac{1}{4}\big[2r_1 r_4 \vartheta_1 + t_1 t_4 \big(\vartheta_3\{\theta_2(\theta_5-\theta_4) + \vartheta_2(\vartheta_4-\vartheta_5)\} \\ &+ \theta_3\{\vartheta_2(\theta_5-\theta_4) + \theta_2(\bar{\theta}_5-\vartheta_4)\}\big)\big]^2, \end{aligned} \tag{4.33}$$

[4] In Ref. [113] Pusey *et al.* investigate the meaning of the quantum state represented by a wavefunction, Ψ. They assume that independently prepared systems have independent physical states, which are objective and independent of the observer. Given these assumptions they find that Ψ must correspond to a real physical state. However, their analysis does not consider the appearance of such a state from the perspective of an observer with a limited amount of information about it.

$$P(n_1^V=1|\boldsymbol{\theta}) = \frac{1}{4}\big[2r_1r_4\theta_1 + t_1t_4\big(\vartheta_2\{\vartheta_3(\theta_4-\theta_5)+\theta_3(\vartheta_4-\vartheta_5)\} + \theta_2\{\theta_3(\theta_5-\theta_4)+\vartheta_3(\vartheta_4-\vartheta_5)\}\big)\big]^2, \tag{4.34}$$

$$P(n_2^H=1|\boldsymbol{\theta}) = \frac{1}{4}\big[2r_1t_4\vartheta_1 + r_4t_1\big(\theta_3\{\vartheta_2(\theta_4-\theta_5)+\theta_2(\vartheta_4-\vartheta_5)\} + \vartheta_3\{\theta_2(\theta_4-\theta_5)+\vartheta_2(\vartheta_5-\vartheta_4)\}\big)\big]^2, \tag{4.35}$$

$$P(n_2^V=1|\boldsymbol{\theta}) = \frac{1}{4}\big[2r_1t_4\theta_1 + r_4t_1\big(\vartheta_2\{\vartheta_3(\theta_5-\theta_4)+\theta_3(\vartheta_5-\vartheta_4)\} + \theta_2\{\theta_3(\theta_4-\theta_5)+\vartheta_3(\vartheta_5-\vartheta_4)\}\big)\big]^2, \tag{4.36}$$

$$P(n_3^H=1|\boldsymbol{\theta}) = \frac{1}{4}t_1^2\big[\theta_2(\theta_4+\theta_5)-\vartheta_2(\vartheta_4+\vartheta_5)\big]^2, \tag{4.37}$$

$$P(n_3^V=1|\boldsymbol{\theta}) = \frac{1}{4}t_1^2\big[\vartheta_2(\theta_4+\theta_5)+\theta_2(\vartheta_4+\vartheta_5)\big]^2, \tag{4.38}$$

where $\vartheta_i \equiv \sqrt{1-\theta_i^2}$.

Classically, restricting the measurement space or imposing post-selection does not affect the ability to deduce the past path of an object. This is because, for identical experiments, a certain outcome corresponds to a deterministic chain of events, which can be backtracked to map out the path of the object. If the most optimal measurement is conducted on a classical object, determinism leads to the Fisher information tending towards infinity. As we know, the quantum scenario is different. The irreversible role of post-selection in quantum mechanics can effectively change the quantum state of the system.

Alice will record a normalised probability distribution given by

$$P(n_1^H=1|\boldsymbol{\theta}) = \mathscr{N}\frac{1}{4}\big[2r_1r_4\vartheta_1 + t_1t_4\big(\vartheta_3\{\theta_2(\theta_5-\theta_4)+\vartheta_2(\vartheta_4-\vartheta_5)\} + \theta_3\{\vartheta_2(\theta_5-\theta_4)+\theta_2(\bar{\theta}_5-\vartheta_4)\}\big)\big]^2, \tag{4.39}$$

$$P(n_1^V=1|\boldsymbol{\theta}) = \mathscr{N}\frac{1}{4}\big[2r_1r_4\theta_1 + t_1t_4\big(\vartheta_2\{\vartheta_3(\theta_4-\theta_5)+\theta_3(\vartheta_4-\vartheta_5)\} + \theta_2\{\theta_3(\theta_5-\theta_4)+\vartheta_3(\vartheta_4-\vartheta_5)\}\big)\big]^2, \tag{4.40}$$

where $\mathscr{N} \equiv \big[P(n_1^H=1|\boldsymbol{\theta})+P(n_1^V=1|\boldsymbol{\theta})\big]^{-1}$.

How does Alice perceive the past of the quantum particle? We can, again, answer this question by calculating the Fisher information that Alice obtains:

$$\lim_{\boldsymbol{\theta}\to 0} F(\theta_1) = 4, \tag{4.41}$$

$$\lim_{\boldsymbol{\theta}\to 0} F(\theta_2) = 0, \tag{4.42}$$

$$\lim_{\boldsymbol{\theta}\to 0} F(\theta_3) = 0, \tag{4.43}$$

$$\lim_{\boldsymbol{\theta}\to 0} F(\theta_4) = \frac{t_1^2 t_4^2}{r_1^2 r_4^2}, \tag{4.44}$$

$$\lim_{\boldsymbol{\theta}\to 0} F(\theta_5) = \frac{t_1^2 t_4^2}{r_1^2 r_4^2}. \tag{4.45}$$

Let me summarise these results. First, Alice is only aware of the events when she receives a particle. Therefore, every event that is used to produce her probability distribution is one where Alice received a particle. Second, Alice receives a quantity of Fisher information of θ_1 equal to the free-space evolution value (see Chapter 3.5.1). This supports the standard picture of the arrangement, that the particle must have passed through the first tagging interaction in order to have reached Alice. The standard picture is also supported by the vanishing Fisher information quantities obtained by Alice with respect to θ_2 and θ_3. Third, we find the surprising result that Alice obtains a finite amount of Fisher information with respect to θ_4 and θ_5, which are placed within the inner MZI of the NMZI device. From Alice's perspective she receives a number of particles from which she can decode information from the continuous path through the first tagging interaction, as well as from the discontinuous path containing the inner MZI.

We have based our definition of the past presence of an object on the information an experimentalist obtains through repeated experiments. Hence, from Alice's perspective the quantum particles she receives have had a past presence at the first tagging interaction as well as within the inner MZI. We also note that an external observer, Oliver, with access to all the output states would interpret the past of the quantum particles differently. We can therefore draw the remarkable conclusion that, from our Fisher information perspective, the perception of the past of a quantum particle is subjectively dependent on the measurements which an observer conducts on it. This is illustrated in Fig. 4.4, which shows where Fisher information can be transmitted to Alice and Oliver, respectively. Note that Fig. 4.4(a) supports the conclusions presented in Fig. 1.12(b) without relying on the TSVF.

In the specific example of the NMZI, Alice obtains particles that have had a discontinuous path through the device. Indeed, in the limit of small $\boldsymbol{\theta}$, the Schrödinger picture shows vanishing amplitudes of wavefunction propagating from BS_3 to BS_4. The mere possibility of propagation in this direction (in the scenario of finite disturbances within the inner MZI) is enough to allow Alice to decode information from the inner part of the device. The non-classical part of this is the discontinuity of this spatial location from the rest of the paths that Alice obtains information from.

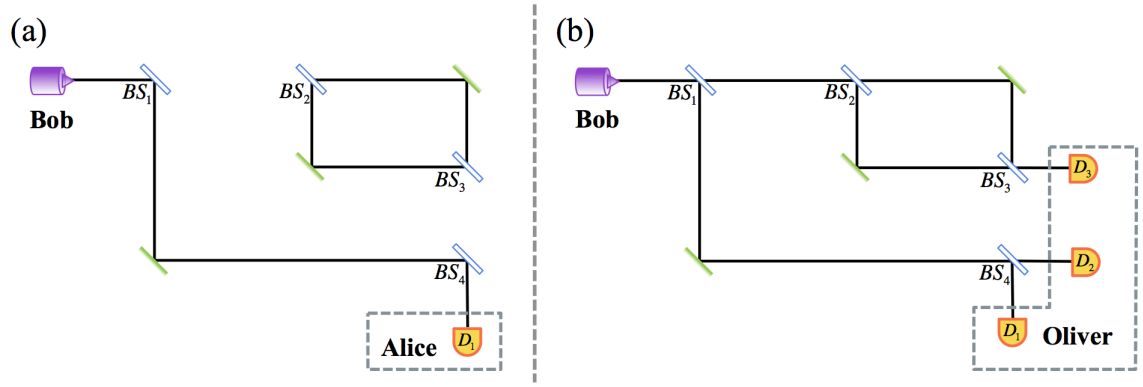


Figure 4.4 The past path of quantum particle (solid line) in an NMZI as experienced by (a) Alice, and (b) Oliver.

6 Conclusion

We have now extended the Fisher information measure for inter-measurement quantum presence. In the previous chapter we saw how the classical Fisher information could be used to analyse inter-measurement paths of particles through the devices used in CFC protocols. In order to account for a more generic scenario, we studied tagging interactions in general quantum circuits subject to generalised measurements (positive-operator valued measures). The formula for the quantum Fisher information obtained from these generic experiments takes a simple and beautiful form. It is the function of the integrated probability density of occupying the two levels of the tagging prior to the interaction. However, we also found that the complex phase relation between the pre-tagging quantum state and the elements of the tagging interaction can have a destructive impact on the quantum Fisher information obtained in an experiment. It is possible for the wavefunction, in the Schrödinger picture, to pass through a tagging interaction in such a way that the tagging parameter is "encoded" in the global phase, leaving no measurable impact at the output. However, as any propagating quantum particle will interact continuously with the surroundings, a realistic scenario is bound to include tagging interactions where the phase relation is not destructive.

In Chapter 1.5 we saw how previous attempts to scrutinise the NMZI device have been based on a weak trace formalism [39]. This analysis hinges upon the TSVF interpretation of quantum mechanics and considers the physical reality of a quantum particle to be represented by the overlapping paths of the forward and backward evolving wavefunction from a pre- and post-selected experiment. In this chapter we have seen that an alternative interpretation of a quantum object (particle) can be based on the actual information which a specific experimental observation of such an object (particle) yields. This information should be the same within all interpretations of quantum mechanics. Hence, if one assumes that our reality

is not described by a non-local hidden variable theory, the analysis of this chapter should be interpretation-independent. On the other hand, if one does assume a quantum theory, such as the de Broglie-Bohm theory (in which point-like particles follow deterministic and well-defined paths), it would rather be a hidden variable (the quantum potential in the case of the de Broglie-Bohm theory) which we perceive as the quantum object.

I think that the otherwise abstract concept of a quantum wavefunction at a specific spatial location, ψ_i, is more easily interpreted within the new Fisher information framework. The quantity $|\psi_k|^2 + |\psi_{k'}|^2$ simply represents the amount of information that *could* be encoded by an interaction acting between level k and k', and the phase, $\arg(\psi_k^c) - \arg(\psi_{k'}^c)$, determines how much information that *will* be encoded by the specific interaction.

An important result of this chapter is the possibility of using the quantum Fisher information measure to investigate several past locations of the same quantum particle. When this is done by an observer that has access to all the output ports of the circuit, the quantum paths appear to follow the well-defined and continuous paths of the wavefunction in the standard Schrödinger picture. However, when post-selection is imposed, the results are different. The paths can then be discontinuous, suggesting that the perception of a quantum particle's inter-measurement past is subjectively dependent on the measurement of the observer.

Chapter 5

Conclusion

In this thesis I have developed a protocol for counterfactual communication, and a methodology for investigating inter-measurement presence of quantum particles. Below I summarise my findings and provide suggestions for future work.

1 Summary

In Chapter 1 I briefly outlined the current state of quantum mechanics. I started by presenting the postulates of quantum mechanics, the toolkit used in order to make accurate predictions of quantum experiments. I gave a brief description of the Copenhagen interpretation of quantum mechanics, the first attempt to interpret the behaviour described by the postulates. We saw that this interpretation does not provide a satisfactory explanation of the physical reality behind the observations of quantum experiments. In particular, the Copenhagen interpretation fails to provide a way of considering quantum states between measurements. I suggested that the development of a better understanding of quantum mechanics relies on the development of new quantum phenomena, which can provide the pieces needed to build the puzzle of the physical reality beyond the Copenhagen interpretation. I presented various results from quantum mechanics and information theory used in this thesis.

We saw how the evolution of quantum states through discrete interferometers can be modelled by the operation of unitary matrices on vectors in a Hilbert space. Examples included polarization rotations, beam-splitters and phase-shifters. We studied how such interferometers can be used for parameter estimation of, for example, phase-shifts. The concept of parameter estimation was generalised with a presentation of the maximum likelihood method, which allows for the estimation of an unknown parameter from the statistics of the measurement outcomes of experiments.

One of the main objectives of this thesis was to develop a framework for discussing the spatial inter-measurement presence of quantum states in interferometers. In Chapter 1 I summarised the work done by previous authors on the interpretation of inter-measurement quantum states within the time-symmetric two-state vector formalism (a niche interpretation of quantum mechanics). The previously used measure of the past of quantum particles relies on the evaluation of weak values, which has been subject to intense discussion and criticism.

I provided a study of the quantum Zeno effect, which allows for the "freezing" of the evolution of a quantum state by repeated measurements of it. I first presented the general theory behind the effect, followed by an example based on a concatenated Mach-Zehnder interferometer. I approached the effect from a slightly different angle to previous works, showing that it is the deprivation of quantum self-interference, rather than measurements, that allows for the manifestations of the quantum Zeno effect.

In the last part of Chapter 1 I introduced the concept of counterfactual quantum phenomena via the specific example of interaction-free measurements. Such measurements have received significant attention owing to their ability to determine the presence of an object without interacting with it.

Chapter 2 contained a detailed study of counterfactual communication, that is communication in which a transmitter can send a message to a receiver without the receiver obtaining any particles that have interacted with the transmitter. I first presented the workings of a concatenated Mach-Zehnder interferometer, the basic structure of counterfactual communication protocols. I then outlined the first suggested counterfactual communication protocol, the type I protocol of Salih *et al.* [39]. We saw how a combined use of nested concatenated Mach-Zehnder interferometers and the quantum Zeno effect seems to allow for information to be transmitted from a transmitter to a receiver, without any particles crossing the transmission line between them. However, we also saw that this scheme would require experimentally unrealistic numbers of optical components, and that a weak trace analysis of the protocol suggests that it actually is not counterfactual.

At this point I presented my own contribution to the study of counterfactual communication, the type II protocol, also referred to as "quantum telepathy" [46]. This protocol uses a different definition of counterfactual communication, where particles are allowed to propagate in the opposite direction to the message (that is, from receiver to transmitter). I developed the protocol within the framework of second quantisation, with which we could see how true counterfactual communication can be realised for both a 0-bit and a 1-bit process. In order to provide a more realistic evaluation, including wavefunctions of finite width, I presented solutions to the time-dependent Schrödinger equation subject to a potential that implements the type II protocol. This analysis made it evident that the protocol is robust

against disturbances throughout the evolution of the protocol particles. We saw that even if the number of optical components is reduced, such that the failure probability of individual protocol processes is high, it is possible to employ an error correction scheme in which each logical bit is decoded from several processes. Using this scheme it is possible to heavily suppress the bit error rate whilst keeping the counterfactual violation rate sufficiently small. In the final part of Chapter 2 I discussed the ontological implications of type II counterfactual communication. We saw a beautiful manifestation of interaction-free quantum non-locality by considering how the local changes of the single-particle Hamiltonian at the transmitter location affects the dynamics of the quantum state in Alice's laboratory, without the exchange of any force-carrying particles.

As counterfactual communication puts constraints on where the protocol particles are allowed to be present, between initialisation and measurement, a proper analysis of such phenomena requires a methodology for the discussion of inter-measurement presence and the departure from the Copenhagen interpretation. Such a methodology was presented in Chapter 3. First, I outlined two measures from information theory: the Shannon mutual information, which describes the suitability of a given experiment for the estimation of an unknown parameter, and the Fisher information, which quantifies the information gain of a specific unknown parameter in an experiment. I showed that, for the specific non-polarizing interferometers studied in this thesis, the Fisher information of a disturbing polarization interaction set by an unknown parameter is proportional to the integrated probability density at the position of the interaction. Any realistic quantum evolution will naturally involve such weak disturbing interactions. As the Fisher information represents the actual information gain at the output of an experiment, I argued that the past presence of a quantum particle can be measured by evaluating the Fisher information of any such interactions which the wavefunction has propagated through.

During recent years, the evolution of quantum states in nested Mach-Zehnder interferometers has been under intense debate. Chapter 3 contained a warranted analysis of the Shannon mutual information and the Fisher information of interactions in these structures. I then provided a thorough analysis of counterfactuality in the previously suggested type I protocols, which are based on nested Mach-Zehnder interferometers. Via a theoretical analysis of a reduced post-selected protocol we saw how the Fisher information measure can be used to show that type I counterfactuality is violated. This finding was supported by a numerical simulation of a wavepacket propagating through a nested Mach-Zehnder interferometer. A numerical method was also used to show that the original protocol for type I counterfactual communication (Ref. [39]) is not truly counterfactual unless the quantum channels used are absolutely perfect. In the presence of weak disturbances this protocol is many orders of

magnitude less counterfactual than direct communication. Finally, my protocol for type II counterfactuality was analysed with a probability measure. This protocol satisfies the type II definition of counterfactuality even in the presence of weak disturbances.

In Chapter 4 I extended the Fisher information measure of inter-measurement particle presence to more general scenarios. First, I discussed the classical concept of an object from an information-theoretical perspective. I then pivoted towards a quantum scenario, suggesting how information can be used to discuss quantum objects (or particles). Instead of considering the non-polarizing interferometers of Chapter 3, I provided a calculation of the quantum Fisher information of a general quantum circuit subject to an unknown two-level interaction. We saw how this information quantity was composed of two terms: the first one being proportional to the integrated probability density at the loci of the interaction; and the second one being a destructive term set by the phases of the pre-interaction wavefunction components and the phase shifts induced by the interaction. I argued that as long as the destructive term is zero (which is the case for most studied interferometers), the quantum Fisher information provides an excellent measure of inter-measurement presence.

I also extended the inter-measurement measure to allow for the joint investigation of the past presence at several locations of the quantum circuit. I used this to revise the study of the nested Mach-Zehnder interferometer. We saw that when all measurement outcomes were taken into account in the analysis, the quantum Fisher measure confirmed the "common sense" picture, that the quantum particle had propagated through continuous paths through the device. However, if an experimentalist only has access to a limited number of output ports of the device, the conclusions can be different. In this case the experimentalist can detect particles that appear to have had discontinuous paths through the device. From this I drew the remarkable conclusion that the inter-measurement presence of a particle is subjectively dependent on the measurement of the observer.

2 Further Work

There are numerous opportunities to produce interesting results by extending the work presented in this thesis. Below I give some examples.

The publication of my type II counterfactual communication protocol (Ref. [46]) received significant attention. I am currently working on the experimental realisation of the protocol together with Philip Walther and Irati Alonso Calafell at the University of Vienna. The experimental protocol is based on a state-of-the art nanophotonic processor. Our preliminary data suggest that this ultrastable platform, together with the error correction scheme presented in Chapter 2.4.5, will allow for the realisation of counterfactual communication with bit error

and counterfactual violation rates reduced to less than 2 %. This work would be the first ever true realisation of interaction-free communication.

The Fisher information measure of type I counterfactual violations (published in Ref. [47]) has been highly successful in providing a new way to discuss particle presence. However, the extension of this theory, presented in Chapter 4, is far broader. I intend to summarise the results in a paper, giving a general overview of how to use information theory to discuss inter-measurement particle presence. In doing this it would make sense to include a full analysis of the quantum Fisher information matrix (see, for example, Ref. [103]), which provides a bound on the covariance between all the unknown parameters considered. This thesis intuits that such an analysis could allow for the discussion of not only the past presence at specific locations, but also of the inter-measurement directional propagation of specific components of a quantum particle.

In private discussions with Avshalom Elitzur and Lev Vaidman I have received suggestions for interesting quantum experiments to study with the Fisher information measure. For example, in Ref. [98] Aharanov *et al.* present a surprising post-selected experiment in which a particle is present, disappears and finally re-appears. This experiment is explained within the framework of the two-state vector formalism, and it is possible that the operational toolkit of the Fisher information measure would allow for a less controversial discussion of the result. Another example is the paper by Alonso *et al*. [114] where dove prisms are introduced within a nested Mach-Zehnder interferometer. The previous weak trace analysis of Danan *et al.* within the two-state vector formalism fails to describe the findings of Alonso [115, 116]. Again it is possible that the Fisher information measure can provide an interpretation-independent solution to this problem.

Finally, this thesis has concerned itself with the development of a better understanding of the foundations of quantum theory. However, many of the tools developed in relation to the Fisher information measure are highly relevant to quantum parameter estimation in general. I am working with Aleksander Lasek and Hugo Lepage on a paper on two-qubit quantum gates in surface acoustic wave systems. The Fisher information measure has proved successful in establishing the experimental viability of certain entangling operations. My hope is to extend this result in order to provide a better framework for the experimental evaluation of quantum computer components.

Bibliography

[1] Chris J Isham. *Lectures on quantum theory: Mathematical and structural foundations*. Allied Publishers, 2001.

[2] Richard I G Hughes. *The structure and interpretation of quantum mechanics*. Harvard University Press, 1992.

[3] B H Bransden and C J Joachain. *Quantum Mechanics (2nd edition)*. Prentice Hall, 2000.

[4] Wojciech Hubert Zurek. Quantum origin of quantum jumps: Breaking of unitary symmetry induced by information transfer in the transition from quantum to classical. *Phys. Rev. A*, 76:052110, Nov 2007.

[5] Werner Heisenberg. *The physical principles of the quantum theory*. Courier Corporation, 1949.

[6] Jagdish Mehra and Helmut Rechenberg. *The Historical Development of Quantum Theory*. Springer-Verlag, 2001.

[7] David Deutsch. Quantum mechanics near closed timelike lines. *Physical Review D*, 44(10):3197, 1991.

[8] Seth Lloyd, Lorenzo Maccone, Raul Garcia-Patron, Vittorio Giovannetti, Yutaka Shikano, Stefano Pirandola, Lee A Rozema, Ardavan Darabi, Yasaman Soudagar, Lynden K Shalm, et al. Closed timelike curves via postselection: theory and experimental test of consistency. *Physical Review Letters*, 106(4):040403, 2011.

[9] Martin Ringbauer, Matthew A Broome, Casey R Myers, Andrew G White, and Timothy C Ralph. Experimental simulation of closed timelike curves. *Nature communications*, 5:4145, 2014.

[10] Asher Peres. Karl popper and the copenhagen interpretation. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 33(1):23–34, 2002.

[11] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Quantum mechanics: non-relativistic theory*, volume 3. Elsevier, 2013.

[12] C. H. W. Barnes. *Part III Physics Lecture Course: Quantum Information*. University of Cambridge, Cambridge, UK, 2018.

[13] N David Mermin. Could Feynman have said this. *Physics Today*, 57(5):10, 2004.

[14] David Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Phys. Rev.*, 85:166–179, Jan 1952.

[15] David Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. ii. *Phys. Rev.*, 85:180–193, Jan 1952.

[16] Hugh Everett III. *The theory of the universal wave function*. Princeton University, 1973.

[17] Lev Vaidman. *Many-Worlds Interpretation of Quantum Mechanics*. Metaphysics Research Lab, Stanford University, fall 2016 edition, 2002.

[18] G C Ghirardi, A Rimini, and T Weber. A model for a unified quantum description of macroscopic and microscopic systems. pages 223–232, 1985.

[19] G. C. Ghirardi, A. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems. *Phys. Rev. D*, 34:470–491, Jul 1986.

[20] Peter J Lewis. *Quantum ontology: a guide to the metaphysics of quantum mechanics*. Oxford University Press, 2016.

[21] Claude Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, July 1948.

[22] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[23] Benjamin Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, Apr 1995.

[24] George Markowsky. Information theory. *Encyclopædia Britannica*, June 2017.

[25] Roger Penrose. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press, Inc., New York, NY, USA, 1st edition, 1994.

[26] Sarah Collins. Researchers chart the 'secret' movement of quantum particles. *University fo Cambridge - Research (online)*, Dec 2017.

[27] India Ashok. Scientists just discovered how to track the secret movement of elusive unobserved quantum particles. *International Business Times*, Dec 2017.

[28] Ludwig Zehnder. *Ein neuer interferenzrefraktor*. Springer, 1891.

[29] Ludwig Mach. Ueber einen interferenzrefraktor. *Zeitschrift für Instrumentenkunde*, 12(3):89, 1892.

[30] Thomas B. Bahder and Paul A. Lopata. Fidelity of quantum interferometers. *Phys. Rev. A*, 74:051801, Nov 2006.

[31] D. S. Simon, A. V. Sergienko, and T. B. Bahder. Dispersion and fidelity in quantum interferometry. *Phys. Rev. A*, 78:053829, Nov 2008.

[32] Thomas B. Bahder. Phase estimation with nonunitary interferometers: Information as a metric. *Phys. Rev. A*, 83:053601, May 2011.

[33] James O Berger. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.

[34] Subhash Sharma. *Applied Multivariate Techniques*. Wiley, 1 edition, 1996.

[35] J. L. Lions F. Cucker, P. G. Ciarlet. *Handbook of numerical analysis - Foundations of computational mathematics*, volume V11. Elsevier Science & Technology, 2003.

[36] L. Vaidman. Past of a quantum particle. *Phys. Rev. A*, 87:052104, May 2013.

[37] Zheng-Hong Li, M. Al-Amri, and M. Suhail Zubairy. Comment on "past of a quantum particle". *Phys. Rev. A*, 88:046102, Oct 2013.

[38] L. Vaidman. Reply to "comment on 'past of a quantum particle' ". *Phys. Rev. A*, 88:046103, Oct 2013.

[39] Hatim Salih, Zheng-Hong Li, M. Al-Amri, and M. Suhail Zubairy. Protocol for direct counterfactual quantum communication. *Phys. Rev. Lett.*, 110:170502, Apr 2013.

[40] Lev Vaidman. Comment on "protocol for direct counterfactual quantum communication". *Phys. Rev. Lett.*, 112:208901, May 2014.

[41] Hatim Salih, Zheng-Hong Li, M. Al-Amri, and M. Suhail Zubairy. Salih *et al.* reply:. *Phys. Rev. Lett.*, 112:208902, May 2014.

[42] Zheng-Hong Li, M. Al-Amri, and M. Suhail Zubairy. Direct counterfactual transmission of a quantum state. *Phys. Rev. A*, 92:052315, Nov 2015.

[43] L. Vaidman. Comment on "direct counterfactual transmission of a quantum state". *Phys. Rev. A*, 93:066301, Jun 2016.

[44] Zheng-Hong Li, M. Al-Amri, and M. Suhail Zubairy. Reply to "comment on 'direct counterfactual transmission of a quantum state"'. *Phys. Rev. A*, 93:066302, Jun 2016.

[45] Robert B. Griffiths. Particle path through a nested mach-zehnder interferometer. *Phys. Rev. A*, 94:032115, Sep 2016.

[46] D. R. M. Arvidsson-Shukur and C. H. W. Barnes. Quantum counterfactual communication without a weak trace. *Phys. Rev. A*, 94:062303, Dec 2016.

[47] D. R. M. Arvidsson-Shukur, A. N. O. Gottfries, and C. H. W. Barnes. Evaluation of counterfactuality in counterfactual communication protocols. *Phys. Rev. A*, 96:062316, Dec 2017.

[48] Yakir Aharonov and Lev Vaidman. *The two-state vector formalism: an updated review*. Springer, 2008.

[49] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan. The sense in which a "weak measurement" of a spin-½ particle's spin component yields a value 100. *Phys. Rev. D*, 40:2112–2117, Sep 1989.

[50] Yakir Aharonov, David Z. Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin- *1/2* particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351–1354, Apr 1988.

[51] Onur Hosten and Paul Kwiat. Observation of the spin hall effect of light via weak measurements. *Science*, 319(5864):787–790, 2008.

[52] Kevin Lyons, Justin Dressel, Andrew N. Jordan, John C. Howell, and Paul G. Kwiat. Power-recycled weak-value-based metrology. *Phys. Rev. Lett.*, 114:170801, Apr 2015.

[53] Omar S. Magaña Loaiza, Mohammad Mirhosseini, Brandon Rodenburg, and Robert W. Boyd. Amplification of angular rotations using weak measurements. *Phys. Rev. Lett.*, 112:200401, May 2014.

[54] P. Ben Dixon, David J. Starling, Andrew N. Jordan, and John C. Howell. Ultrasensitive beam deflection measurement via interferometric weak value amplification. *Phys. Rev. Lett.*, 102:173601, Apr 2009.

[55] B. Misra and E. C. G. Sudarshan. The zeno's paradox in quantum theory. *Journal of Mathematical Physics*, 18(4), 1977.

[56] Aristotle. Physics. page 239b Bekker, 350BC.

[57] Alan Mathison Turing, Robin O Gandy, Charles Edmund Michael Yates, and Solomon Feferman. *Mathematical logic*. Elsevier Science, 2001.

[58] A. Degasperis, L. Fonda, and G. C. Ghirardi. Does the lifetime of an unstable system depend on the measuring apparatus? *Il Nuovo Cimento A (1965-1970)*, 21(3):471–484, 1974.

[59] Avshalom C. Elitzur and Lev Vaidman. Quantum mechanical interaction-free measurements. *Foundations of Physics*, 23(7):987–997, 1993.

[60] Andrew G. White, Jay R. Mitchell, Olaf Nairz, and Paul G. Kwiat. "interaction-free" imaging. *Phys. Rev. A*, 58:605–613, Jul 1998.

[61] Serge Massar, Graeme Mitchison, and Stefano Pironio. Minimal absorption measurements. *Phys. Rev. A*, 64:062303, Nov 2001.

[62] L. Chirolli, E. Strambini, V. Giovannetti, F. Taddei, V. Piazza, R. Fazio, F. Beltram, and G. Burkard. Electronic implementations of interaction-free measurements. *Phys. Rev. B*, 82:045403, Jul 2010.

[63] John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[64] Lucien Hardy. Quantum mechanics, local realistic theories, and lorentz-invariant realistic theories. *Phys. Rev. Lett.*, 68:2981–2984, May 1992.

[65] Stephen Wiesner. Conjugate coding. *SIGACT News*, 15(1):78–88, January 1983.

[66] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, 1982.

[67] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. volume 175, pages 175–179, 1984.

[68] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[69] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and epr channels. 1993.

[70] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409, 2001.

[71] Peter W. Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, Jul 2000.

[72] Zhang Quan and Tang Chaojing. Simple proof of the unconditional security of the bennett 1992 quantum key distribution protocol. *Phys. Rev. A*, 65:062301, May 2002.

[73] Dagomir Kaszlikowski, D. K. L. Oi, Matthias Christandl, Kelken Chang, Artur Ekert, L. C. Kwek, and C. H. Oh. Quantum cryptography based on qutrit bell inequalities. *Phys. Rev. A*, 67:012310, Jan 2003.

[74] Tae-Gon Noh. Counterfactual quantum cryptography. *Phys. Rev. Lett.*, 103:230501, Dec 2009.

[75] Paul Kwiat, Harald Weinfurter, Thomas Herzog, Anton Zeilinger, and Mark A. Kasevich. Interaction-free measurement. *Phys. Rev. Lett.*, 74:4763–4766, Jun 1995.

[76] P. G. Kwiat, A. G. White, J. R. Mitchell, O. Nairz, G. Weihs, H. Weinfurter, and A. Zeilinger. High-efficiency quantum interrogation measurements via the quantum zeno effect. *Phys. Rev. Lett.*, 83:4725–4728, Dec 1999.

[77] D. R. M Arvidsson-Shukur. Quantum information - first year phd report in physics. *University of Cambridge*, 2015.

[78] Onur Hosten, Matthew T. Rakher, Julio T. Barreiro, Nicholas A. Peters, and Paul G. Kwiat. Counterfactual quantum computation through quantum interrogation. *Nature*, 439, February 2006.

[79] L Vaidman. Counterfactuality of 'counterfactual' communication. *Journal of Physics A: Mathematical and Theoretical*, 48(46):465303, 2015.

[80] D. R. M. Arvidsson-Shukur, H. V. Lepage, E. T. Owen, T. Ferrus, and C. H. W. Barnes. Protocol for fermionic positive-operator-valued measures. *Phys. Rev. A*, 96:052305, Nov 2017.

[81] S. E. Ahnert and M. C. Payne. All possible bipartite positive-operator-value measurements of two-photon polarization states. *Phys. Rev. A*, 73:022333, Feb 2006.

[82] C. H. W. Barnes, J. M. Shilton, and A. M. Robinson. Quantum computation using electrons trapped by surface acoustic waves. *Phys. Rev. B*, 62:8410–8419, Sep 2000.

[83] A. A. Michelson and E. W. Morley. On the relative motion of the earth and the luminiferous ether. *American Journal of Science*, 32(203):333–345, 1887.

[84] B. P. Abbott and et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.

[85] P. Hariharan. *Optical Interferometry*. Academic Press, Cambridge, Massachusetts, USA, 2nd edition, 2003.

[86] G Bertocchi, O Alibart, D B Ostrowsky, S Tanzilli, and P Baldi. Single-photon sagnac interferometer. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 39(5):1011, 2006.

[87] Emanuele Berti, Jaime Cardoso, Vitor Cardoso, and Marco Cavaglià. Matched filtering and parameter estimation of ringdown waveforms. *Phys. Rev. D*, 76:104044, Nov 2007.

[88] John K. Stockton, Xinan Wu, and Mark A. Kasevich. Bayesian estimation of differential interferometer phase. *Phys. Rev. A*, 76:033613, Sep 2007.

[89] R. Demkowicz-Dobrzanski, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley. Quantum phase estimation with lossy interferometers. *Phys. Rev. A*, 80:013825, Jul 2009.

[90] Hugo Cable and Gabriel A. Durkin. Parameter estimation with entangled photons produced by parametric down-conversion. *Phys. Rev. Lett.*, 105:013603, Jul 2010.

[91] S. Vitale and M. Zanolin. Parameter estimation from gravitational waves generated by nonspinning binary black holes with laser interferometers: Beyond the fisher information. *Phys. Rev. D*, 82:124065, Dec 2010.

[92] Lee Rozema, Ardavan Darabi, Dylan Mahler, Alex Hayat, Yasaman Soudagar, and Aephraim M. Steinberg. Direct violation of heisenberg's precision limit by weak measurements. In *Frontiers in Optics 2012/Laser Science XXVIII*, page FW4J.4. Optical Society of America, 2012.

[93] Lee A. Rozema, Ardavan Darabi, Dylan H. Mahler, Alex Hayat, Yasaman Soudagar, and Aephraim M. Steinberg. Violation of heisenberg's measurement-disturbance relationship by weak measurements. *Phys. Rev. Lett.*, 109:100404, Sep 2012.

[94] D. H. Mahler, Lee A. Rozema, Ardavan Darabi, Christopher Ferrie, Robin Blume-Kohout, and A. M. Steinberg. Adaptive quantum state tomography improves accuracy quadratically. *Phys. Rev. Lett.*, 111:183601, Oct 2013.

[95] Z. Hradil. Estimation of counted quantum phase. *Phys. Rev. A*, 51:1870–1873, Mar 1995.

[96] Ruth E. Kastner. The transactional interpretation, counterfactuals, and weak values in quantum theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 39(4):806 – 818, 2008.

[97] Bengt E. Y. Svensson. What is a quantum-mechanical "weak value" the value of? *Foundations of Physics*, 43(10):1193–1205, Oct 2013.

[98] Yakir Aharonov, Eliahu Cohen, Ariel Landau, and Avshalom C Elitzur. The case of the disappearing (and re-appearing) particle. *Scientific reports*, 7(1):531, 2017.

[99] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons Inc., Hoboken, New Jersey, USA, 2nd edition, 2006.

[100] B Roy Frieden. *Science from Fisher information: a unification*. Cambridge University Press, 2004.

[101] Samuel L. Braunstein and Carlton M. Caves. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.*, 72:3439–3443, May 1994.

[102] Akio Fujiwara and Hiroshi Nagaoka. Quantum fisher metric and estimation for pure state models. *Physics Letters A*, 201(2):119 – 124, 1995.

[103] Dénes Petz. *Quantum information theory and quantum statistics*. Springer Science & Business Media, 2007.

[104] D. Petz and C. Ghinea. *Introduction to quantum Fisher information*, pages 261–281. World Scientific, 2011.

[105] A. R. Marlow. *Mathematical Foundations of Quantum Theory*. Academic press, Inc. (London) Ltd., London, UK, 1st edition, 1978.

[106] Christopher Ferrie and Joshua Combes. How the result of a single coin toss can turn out to be 100 heads. *Phys. Rev. Lett.*, 113:120404, Sep 2014.

[107] Matthew F. Pusey. Anomalous weak values are proofs of contextuality. *Phys. Rev. Lett.*, 113:200401, Nov 2014.

[108] Lev Vaidman, Alon Ben-Israel, Jan Dziewior, Lukas Knips, Mira Weißl, Jasmin Meinecke, Christian Schwemmer, Ran Ber, and Harald Weinfurter. Weak value beyond conditional expectation value of the pointer readings. *Phys. Rev. A*, 96:032114, Sep 2017.

[109] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in Fortran*, volume 1. Press Syndicate of the University of Cambridge, Cambridge, UK, 2nd edition, 1992.

[110] David H Sanford. *Determinates vs. determinables*. Metaphysics Research Lab, Stanford University, 2008.

[111] Jessica Wilson. *Determinables and Determinates*. Metaphysics Research Lab, Stanford University, spring 2017 edition, 2017.

[112] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho. Quantum speed limit for physical processes. *Phys. Rev. Lett.*, 110:050402, Jan 2013.

[113] Matthew F Pusey, Jonathan Barrett, and Terry Rudolph. On the reality of the quantum state. *Nature Physics*, 8(6):475, 2012.

[114] Miguel A. Alonso and Andrew N. Jordan. Can a dove prism change the past of a single photon? *Quantum Studies: Mathematics and Foundations*, 2(3):255–261, Sep 2015.

[115] A. Danan, D. Farfurnik, S. Bar-Ad, and L. Vaidman. Asking photons where they have been. *Phys. Rev. Lett.*, 111:240402, Dec 2013.

[116] Lev Vaidman and Izumi Tsutsui. When photons are lying about where they have been. *arXiv preprint arXiv:1806.00695*, 2018.

[117] Abraham Goldberg, Harry M. Schey, and Judah L. Schwartz. Computer-generated motion pictures of one-dimensional quantum-mechanical transmission and reflection phenomena. *American Journal of Physics*, 35(3):177–186, 1967.

[118] A. Askar and A. S. Cakmak. Explicit integration method for the time-dependent schrödinger equation for collision problems. *The Journal of Chemical Physics*, 68(6):2794–2798, 1978.

[119] Jon J. V. Maestri, Rubin H. Landau, and Manuel J. Paez. Two-particle Schrödinger equation animations of wave packet–wave packet scattering. *American Journal of Physics*, 68(12):1113–1119, 2000.

[120] E. T. Owen, M. C. Dean, and C. H. W. Barnes. Coherent-state spin qubits in the presence of spin-orbit coupling. *Phys. Rev. A*, 89:032305, Mar 2014.

[121] S. E. Kronin. *Computational Physics*. Benjamin Cummings, Reading, UK, 1986.

Appendix A

Numerical Solution to the TDSE

This thesis contains a couple of simulations of the single-particle time-dependent Schrödinger equation (TDSE). It is given by

$$\hat{H}\Psi(\bar{r},t) = i\hbar\partial_t\Psi(\bar{r},t), \tag{A.1}$$

where ∂_t is the partial derivative with respect to time, $\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\bar{r},t)$ is the space and time dependent Hamiltonian of the system, and $\Psi(\bar{r},t)$ is the wavefunction. In this appendix I present an efficient algorithm for numerical solutions to this equation. The method is called the staggered leapfrog algorithm. A detailed outline of it can be found in Refs. [117–119]. In Ref. [80] I have used this powerful algorithm to study positive-operator-valued measures of massive Fermionic particles in realistic potentials of semiconductor heterostructures. It can also be used to study many-particle behaviour, as is done in Ref. [120].

In the staggered leapfrog algorithm, the spatial and temporal domains of the wavefunction are discretised using a finite difference method. The second order derivative of the kinetic term can be expanded as

$$\begin{aligned}\nabla^2\Psi(x,y,z) \approx & \frac{\Psi(x+\Delta x) - 2\Psi(x) + \Psi(x-\Delta x)}{\Delta x^2} \\ & + \frac{\Psi(y+\Delta y) - 2\Psi(y) + \Psi(y-\Delta y)}{\Delta y^2} \\ & + \frac{\Psi(z+\Delta z) - 2\Psi(z) + \Psi(z-\Delta z)}{\Delta z^2}.\end{aligned} \tag{A.2}$$

In order to simplify the discussion we rewrite our wavefunction using a discretised notation:

$$\psi^n_{k,l,m} \equiv \Psi(x = k\Delta x, y = l\Delta y, z = m\Delta z, t = n\Delta t).$$

The time-dependent potential of the Hamiltonian can also be discretised such that: $V^n_{k,l,m} \equiv V(k\Delta x, l\Delta y, m\Delta z, n\Delta t)$. By generalising the methods from Ref. [119] to 3D we can write our Hamiltonian as

$$\hat{H}\psi^n_{k,l,m} = \frac{\hbar^2}{2m}\left[\frac{2\psi^n_{k,l,m} - \psi^n_{k+1,l,m} - \psi^n_{k-1,l,m}}{\Delta x^2} + \frac{2\psi^n_{k,l,m} - \psi^n_{k,l+1,m} - \psi^n_{k,l-1,m}}{\Delta y^2} + \frac{2\psi^n_{k,l,m} - \psi^n_{k,l,m+1} - \psi^n_{k,l,m-1}}{\Delta z^2}\right] + V^n_{k,l,m}\psi^n_{k,l,m} \tag{A.3}$$

This expression can then be used to approximate the time evolution of the wavefunction using the Euler method:

$$\psi^{n+1}_{k,l,m} = \exp(-i\Delta t H/\hbar)\psi^n_{k,l,m} \approx (1 - i\Delta t H/\hbar)\psi^n_{k,l,m}. \tag{A.4}$$

This simple approximation is, however, not stable over several time iterations [121]. The staggered leapfrog algorithm instead approximates the evolution by

$$\psi^{n+1}_{k,l,m} - \psi^{n-1}_{k,l,m} \approx -(2i\Delta t H/\hbar)\psi^n_{k,l,m}. \tag{A.5}$$

Substituting this expression into the TDSE, and using the Hamiltonian from Eq. A.3, we find that

$$\psi^{n+1}_{k,l,m} \approx \psi^{n-1}_{k,l,m} + \left(\frac{-2i\Delta t}{\hbar}\right)V^n_{k,l,m}\psi^n_{k,l,m} + \frac{i\hbar\Delta t}{m}\left[\frac{2\psi^n_{k,l,m} - \psi^n_{k+1,l,m} - \psi^n_{k-1,l,m}}{\Delta x^2} + \frac{2\psi^n_{k,l,m} - \psi^n_{k,l+1,m} - \psi^n_{k,l-1,m}}{\Delta y^2} + \frac{2\psi^n_{k,l,m} - \psi^n_{k,l,m+1} - \psi^n_{k,l,m-1}}{\Delta z^2}\right]. \tag{A.6}$$

This formula can be used to iteratively solve the TDSE. However, by separating the wavefunction into its real and imaginary components a more efficient solution can be obtained. The wave function is decomposed such that $\psi^n_{k,l,m} = u^n_{k,l,m} + iv^{n'}_{k,l,m}$. We can then evaluate the imaginary component at a staggered time in relation to the real component: $n' = n + 1/2$. An advantage of this is that the numerical iterations can be done using only real numbers. Finally, we extend the methods of Refs. [118, 119] further by including a spin-dependent component such that the potential takes the form of

$$V^n_{k,l,m} \to V^n_{k,l,m} + \sum_{i,j=\uparrow,\downarrow} S^n_{k,l,m,(i,j)}, \tag{A.7}$$

where $S^n_{k,l,m,(i,j)}$ are the coefficients of a spin-dependent part of the potential, which acts on the wavefunction of spin j rotating it to spin i. Note that a more rigorous consideration of spin would include the scalar and vector potentials instead. However, for the scenarios considered in this letter the Lorentz term is negligible and the simplification to the spin-dependent dynamics is reasonable (See Refs. [80, 120]).

We now arrive at the staggered leapfrog algorithm, which performs the following iteration:

$$\begin{aligned}
u^{n+1}_{k,l,m,i=(\uparrow,\downarrow)} \approx\ & u^{n-1}_{k,l,m,i} + \left(\frac{2\Delta t}{\hbar}\right) V^n_{k,l,m,i} v^n_{k,l,m,i} \\
& + \frac{\hbar\Delta t}{m}\left[\frac{2v^n_{k,l,m,i} - v^n_{k+1,l,m,i} - v^n_{k-1,l,m,i}}{\Delta x^2}\right. \\
& + \frac{2v^n_{k,l,m,i} - v^n_{k,l+1,m,i} - v^n_{k,l-1,m,i}}{\Delta y^2} \\
& \left. + \frac{2v^n_{k,l,m,i} - v^n_{k,l,m+1,i} - v^n_{k,l,m-1,i}}{\Delta z^2}\right] \\
& + \left(\frac{2\Delta t}{\hbar}\right) \sum_{j=\uparrow,\downarrow} \Im(S^n_{k,l,m,(i,j)}) u^n_{k,l,m,j} \\
& + \left(\frac{2\Delta t}{\hbar}\right) \sum_{j=\uparrow,\downarrow} \Re(S^n_{k,l,m,(i,j)}) v^n_{k,l,m,j},
\end{aligned} \tag{A.8}$$

for the real part and

$$\begin{aligned}
v^{n'+1}_{k,l,m,i=(\uparrow,\downarrow)} \approx\ & v^{n'-1}_{k,l,m,i} - \left(\frac{2\Delta t}{\hbar}\right) V^n_{k,l,m,i} u^n_{k,l,m,i} \\
& - \frac{\hbar\Delta t}{m}\left[\frac{2u^n_{k,l,m,i} - u^n_{k+1,l,m,i} - u^n_{k-1,l,m,i}}{\Delta x^2}\right. \\
& + \frac{2u^n_{k,l,m,i} - u^n_{k,l+1,m,i} - u^n_{k,l-1,m,i}}{\Delta y^2} \\
& \left. + \frac{2u^n_{k,l,m,i} - u^n_{k,l,m+1,i} - u^n_{k,l,m-1,i}}{\Delta z^2}\right] \\
& + \left(\frac{2\Delta t}{\hbar}\right) \sum_{j=\uparrow,\downarrow} \Im(S^n_{k,l,m,(i,j)}) v^n_{k,l,m,j} \\
& - \left(\frac{2\Delta t}{\hbar}\right) \sum_{j=\uparrow,\downarrow} \Re(S^n_{k,l,m,(i,j)}) u^n_{k,l,m,j},
\end{aligned} \tag{A.9}$$

for the imaginary part. These steps allows for the explicit solution to the TDSE of a wavefunction propagating through a general potential. If the greatest eigenvalue of the Hamiltonian is E_{max}, then the time-step has to satisfy $\Delta t < E_{max}/\hbar$ for the solution to be

stable [119]. Moreover, this algorithm provides an excellent way of obtaining stable solutions to the TDSE using GPU-boosted code.

Appendix B

Classical Fisher Information - General Circuits

In Chapter 3 I derived an expression, Eq. 3.21, for the classical Fisher information given some assumptions on the quantum input state and the single-photon interferometer. The assumptions included phase restrictions on the input wavefunction and on the matrix elements of the unitary operations of the quantum circuit. Moreover, the "tagging" mechanism, which models the inevitable disturbances, was defined as a polarization rotation. In this appendix I shall continue to consider single-photon quantum states, but produce a general formula for the classical Fisher information of a two-level interaction in a quantum circuit.

Again, we define an input vector, $\boldsymbol{a}$, now of length n, evolving into an output vector, $\boldsymbol{b}$. The n levels should cover the complete description of the single-photon quantum state in the circuit of interest. For example, for a photon propagating through an optical circuit with m spatial paths (external degrees of freedom) we require that $n = xm$, where x is the size of the internal Hilbert space of the photon. For a standard ($l = \pm 1$) photon, we have $x = 4$ corresponding to the polarization and the internal angular momentum degrees of freedom.

Similarly to Eq. 3.7, the scattering matrix, $\mathscr{S}$, of the quantum circuit can be described by three unitary operations:

$$\mathscr{S}\boldsymbol{a} \equiv \left[\hat{V} \cdot \hat{f}^{(k,k')}(\theta) \cdot \hat{U}\right]\boldsymbol{a} \equiv \boldsymbol{b}^{(k,k')}, \tag{B.1}$$

where $\hat{U}$ and $\hat{V}$ are the unitary operators representing the quantum evolution up to and after the tagging rotation, respectively, and $\hat{f}^{(k,k')}(\theta)$ is the unitary operator that now describes the *single* tagging rotation (set by the parameter θ) between the specific levels of k and k'. We do not impose any restrictions on k and k', but we can, without loss of generality, define them such that $k < k'$. A sketch of the optical circuit of $\mathscr{S}$ is given in Fig. B.1.

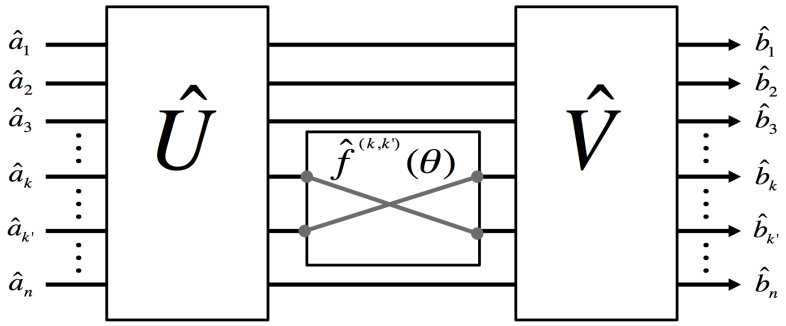


Figure B.1 Sketch of a general optical circuit of the form of Eq. B.1, which is described in the main text.

As in Chapter 3, the wavefunction after the first unitary operation has been applied is given by: $\boldsymbol{c} \equiv \hat{U}\boldsymbol{a}$, with

$$c_i = \sum_{j=1}^{n} U_{i,j} a_j. \tag{B.2}$$

The intermediate wavefunction, $\boldsymbol{c}$, is then subject to the tagging interaction. This results in a rotation between two of the vector levels, k and k'. This time we omit the requirement of Chapter 3, that $k' = k+n$, such that the rotation can be between any levels. The evolution through the tagging interaction is represented by

$$\hat{f}^{(k,k')}_{\phi_a,\phi_b}(\theta) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & e^{i\phi_a} f_{k,k}(\theta) & \cdots & e^{i\phi_b} f_{k,k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-i\phi_b} f_{k',k}(\theta) & \cdots & e^{-i\phi_a} f_{k',k'}(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \tag{B.3}$$

where $\phi_{a,b}$ sets the unitary phases, and we require that $f_{k,k}(\theta) = \sqrt{1-\left[f_{k,k'}(\theta)\right]^2}$, $f_{k',k}(\theta) = -f_{k,k'}(\theta)$ and $f_{k',k'}(\theta) = f_{k,k}(\theta)$, with $f_{k,k}(\theta) : \mathbb{R} \to [-1,1]$.

We express the quantum state after the tagging interaction as $\boldsymbol{d}^{(k,k')}(\theta) \equiv \hat{f}^{(k,k')}(\theta)\boldsymbol{c}$, with

$$d_i^{(k,k')}(\theta) = \sum_{j=1}^{n} f_{i,j}^{(k,k')}(\theta) c_j. \tag{B.4}$$

We note that the only components of $\boldsymbol{d}^{(k,k')}$ that depend on θ are $d_{i=k}^{(k,k')} = e^{i\phi_a} f_{k,k}(\theta)c_k + e^{i\phi_b} f_{k,k'}(\theta)c_{k'}$ and $d_{i=k'}^{(k,k')} = e^{-i\phi_b} f_{k',k}(\theta)c_k + e^{-i\phi_a} f_{k',k'}(\theta)c_{k'}$.

Finally, we can apply the last unitary evolution $\hat{V}$. Following the steps in Chapter 3, we express the output vector as $\boldsymbol{b}^{(k,k')} \equiv \hat{V}\boldsymbol{d}^{(k,k')}(\theta)$, with

$$b_i^{(k,k')}(\theta) \equiv \beta_i^{(k,k')} + b_{i,\theta}^{(k,k')}(\theta) = \sum_{j=1}^{n} V_{i,j} d_j^{(k,k')}(\theta), \tag{B.5}$$

where the θ-dependence of $b_i^{(k,k')}(\theta)$ is encapsulated in $b_{i,\theta}^{(k,k')}(\theta)$, such that

$$\begin{aligned} b_{i,\theta}^{(k,k')}(\theta) \equiv & V_{i,k}\left(e^{i\phi_a} f_{k,k}(\theta)c_k + e^{i\phi_b} f_{k,k'}(\theta)c_{k'}\right) \\ & + V_{i,k'}\left(-e^{-i\phi_b} f_{k,k'}(\theta)c_k + e^{-i\phi_a} f_{k,k}(\theta)c_{k'}\right), \end{aligned}$$

and a corresponding term, independent of θ, is defined as

$$\beta_i^{(k,k')} = \sum_{j\neq k,k'}^{n} V_{i,j} d_j^{(k,k')}(\theta) = \sum_{j\neq k,k'}^{n} V_{i,j} c_j.$$

The probability of measuring the single-photon in the i^{th} output port, is then given by

$$P_i(\theta) = |\beta_i + b_{i,\theta}(\theta)|^2, \tag{B.6}$$

where we, again drop the (k,k')-superscript.

Equation 3.2 is used to calculate the individual Fisher information components:

$$F_i = \frac{1}{\left|\beta_i + b_{i,\theta}(\theta)\right|^2}\left[\frac{\partial}{\partial\theta}\left|\beta_i + b_{i,\theta}(\theta)\right|^2\right]^2 \tag{B.7}$$

As in Chapter 3 the component formula can be simplified by expressing the coefficients as $\beta_i \equiv |\beta_i| e^{i\phi_i}$ and $b_{i,\theta}(\theta) \equiv |b_{i,\theta}(\theta)| e^{i\phi_{i,\theta}}$ and defining $\Phi_{i,\theta}^I \equiv \phi_i - \phi_{i,\theta}$:

$$\begin{aligned} F_i = & \frac{1}{\left|\beta_i\right|^2 + 2\cos\left(\Phi_{i,\theta}^I\right)\left|\beta_i\right|\left|b_{i,\theta}(\theta)\right| + \left|b_{i,\theta}(\theta)\right|^2} \\ & \times \left[\frac{\partial}{\partial\theta}\left(\left|\beta_i\right|^2 + 2\cos\left(\Phi_{i,\theta}^I\right)\left|\beta_i\right|\left|b_{i,\theta}(\theta)\right| + \left|b_{i,\theta}(\theta)\right|^2\right)\right]^2 \\ = & \frac{\left(\cos\left(\Phi_{i,\theta}^I\right)\left|\beta_i\right| + \left|b_{i,\theta}(\theta)\right|\right)^2}{\left|\beta_i\right|^2 + 2\cos\left(\Phi_{i,\theta}^I\right)\left|\beta_i\right|\left|b_{i,\theta}(\theta)\right| + \left|b_{i,\theta}(\theta)\right|^2} \times 4\left[\frac{\partial}{\partial\theta}\left|b_{i,\theta}(\theta)\right|\right]^2, \end{aligned} \tag{B.8}$$

such that

$$F_i \equiv K_{i,\theta}(\theta, \Phi^I_{i,\theta}) \times 4\Big[\frac{\partial}{\partial\theta}\big|b_{i,\theta}(\theta)\big|\Big]^2 \tag{B.9}$$

In order to simplify our expression further, we define $\dot{x} \equiv \partial x/\partial\theta$, as well as $\dot{b}_{i,\theta}(\theta) \equiv \big|\dot{b}_{i,\theta}(\theta)\big|e^{i\phi'_{i,\theta}}$. We also define $\Phi^{II}_{i,\theta} \equiv \phi_{i,\theta} - \phi'_{i,\theta}$. We can then express the total classical Fisher information as

$$\begin{aligned} F(\theta) =& \sum_{i=1}^{n} K_{i,\theta}(\theta, \Phi^I_{i,\theta}) \frac{1}{\big|b_{i,\theta}(\theta)\big|^2}\Big(\frac{\partial}{\partial\theta}\big|b_{i,\theta}(\theta)\big|^2\Big)^2 \\ =& 2\sum_{i=1}^{n} K_{i,\theta}(\theta, \Phi^I_{i,\theta})\big[1+\cos(\Phi^{II}_{i,\theta})\big]\big|\dot{b}_{i,\theta}(\theta)\big|^2. \end{aligned} \tag{B.10}$$

The algebra leading up to Eq. B.10 is straightforward but tedious. Finally, the formula for the classical Fisher information for a general single-photon quantum circuit takes a simple form. The first thing to comment on is the fact that the summation of Eq. B.10 yields the largest possible Fisher information for $\Phi^I_{i,\theta} = \pi m_I$ and $\Phi^{II}_{i,\theta} = 2\pi m_{II}$ with $\{m_I, m_{II}\} \in \mathbb{Z}$. Any other values of Φ^I and Φ^{II} will lead to less constructive summations of the individual Fisher components. It is thus clear that the complex phase of the quantum particle's wavefunction plays an essential role in determining how much information can propagate from the tagging interaction to the output of the interferometer.

If constructive phase relations are satisfied, Eq. B.10 reduces to

$$F(\theta) = 4\sum_{i=1}^{n}\big|\dot{b}_{i,\theta}(\theta)\big|^2. \tag{B.11}$$